

Experimental Democracy –  
Collective Intelligence for a Diverse and Complex World

Felix Gerlsbeck

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy  
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

© 2013  
Felix Gersbeck  
All rights reserved

## Abstract

### Experimental Democracy – Collective Intelligence for a Diverse and Complex World

Felix Gerlsbeck

My dissertation is motivated by the following observation: while we care very much about the *outcomes* of the democratic process, there is widespread uncertainty about *ex ante* how to produce them—and quite often there is also disagreement and uncertainty about what they are in the first place. Consequently, unless we have a definite idea what “better decision-making” might be, it is not obvious which institutional reforms or changes in democratic structures would actually promote it. Democracy is a wide concept, and not all institutional constellations and rules and regulations that can be called democratic function equally well.

In this dissertation therefore I offer a specific model of democracy—“Experimental Democracy”—that unites the view that the quality of decisions matter, with taking into account the circumstances of uncertainty and disagreement that define political problems. On this account, a desirable political mechanism is one that realizes an *experimental method* of policy-making directed at solving problems, such that we can expect it to make progress over time, even though we cannot rule out that it will get things wrong—possibly even frequently. I also show how democracy may best realize such an experimental method, and which particular institutional features of democracy could serve this purpose.

The argument in the dissertation proceeds as follows. In the first part I develop a theory of the justifiability of political authority in the sense outlined above: a theory that is sensitive to the *outcome* concerns that many people share, but recognizes the fundamental disagreement surrounding this question. I establish that instrumental considerations should be of crucial importance when we evaluate political authority. Here I argue against pure proceduralist theories that see the outcome dimension as secondary. However, the facts of disagreement and uncertainty about the ends of politics, as well as concrete policy, do seem to pose a problem for any instrumental justification. In response I outline a *pragmatic* or *experimental* theory of political authority, which focuses precisely on the capacity of a political procedure to solve political problems under uncertainty. Just as in many other fields of inquiry experimentation and adaptation are seen as the adequate responses to uncertainty, I argue, an experimental and adaptive mode of policy-making is the best response to political uncertainty.

In the second part I answer the question which form of democracy would best realize the ideal of experimental policy-making. Subsequently, we should evaluate democratic institutions mainly by their capacity to enable successful experimentation and adaptation. Here, contrary to popular “wisdom of crowds” arguments, I argue that since no single decision procedure can be expected to be reliable across the board, a justified political system may have to employ a plurality of first-order decision-making mechanisms. However, as I show for this to work, these mechanisms must be subject to effective democratic control. The key function of democratic institutions here is that of *feedback*, in order to enable successful adaptation. Finally, I offer some concrete examples how the functional requirements of a successful experimental strategy of policy-making can be institutionally realized within democratic systems.

# Table of Contents

|  |     |
|--|-----|
| Table of Contents  | i   |
| List of Figures and Tables                                   | iii |
| Acknowledgements   | iv  |
| Dedication   | vi  |
| <br>   |     |
| 1. Introduction: What Should a Theory of Democracy Do?       | 1   |
| 1.1 An Icelandic Introduction                                | 1   |
| 1.2 Differentiating Democracy                                | 9   |
| 1.3 The Argument and the Level of Analysis                   | 17  |
| 1.4 The Normative Basis                                      | 20  |
| 1.5 The Argument Coming Up                                   | 22  |
| <br>   |     |
| <i>Part I: A Theory of Experimental Political Legitimacy</i> |     |
| <br>   |     |
| 2. The Value of Democracy                                    | 32  |
| 2.1 Introduction   | 32  |
| 2.2 Justifying Rule; Democratic and Otherwise                | 37  |
| 2.3 What's At Stake?   | 44  |
| 2.4 The Indeterminacy of Pure Proceduralism                  | 47  |
| 2.5 Substantive Values Coming In                             | 53  |
| 2.6 Practical Concerns                                       | 59  |
| 2.7 Intermediate Summary                                     | 64  |
| 2.8 Against Hybrid Principles                                | 65  |
| 2.9 Political Equality as a Residual Value                   | 73  |
| 2.10 Sources of Disagreement                                 | 83  |
| 2.11 Is Politics Really That Difficult?                      | 91  |
| 2.12 Conclusion  | 95  |
| <br>   |     |
| 3. Robust Universal Reliability of Democracy                 | 97  |
| 3.1 Is "Universal Reliability" Possible?                     | 97  |
| 3.2 The Possibility of Rational Choice Under Uncertainty     | 111 |

|  |  |     |
|--|--|-----|
| 3.3  | The Public Interest and the Common Good            | 127 |
| 3.4  | Universal Reliability and How to Argue For It      | 133 |
| 3.5  | Scientists, Not Jurors                             | 150 |
| 3.6  | A Pragmatic Model of Politics                      | 156 |
| 3.7  | Conclusion   | 159 |
| 4.   | Can Experimental Political Authority Be Justified? | 162 |
| 4.1  | Introduction                                       | 162 |
| 4.2  | Complexity, Uncertainty, Difficulty                | 166 |
| 4.3  | What is Experimental Politics?                     | 178 |
| 4.4  | Can We Experiment With Values?                     | 183 |
| 4.5  | Should We Experiment With Values?                  | 195 |
| 4.6  | Conclusion: What Does This Spell for Democracy?    | 212 |
| <br><i>Part II: A Theory of Experimental Democracy</i> |  |     |
| 5.   | Models of Democracy                                | 218 |
| 5.1  | Introduction                                       | 218 |
| 5.2  | Aggregative/Majoritarian Models of Democracy       | 227 |
| 5.3  | Deliberative Models of Democracy                   | 243 |
| 5.4  | Experimental <i>Democratic</i> Legitimacy          | 262 |
| 5.5  | The Control Model of Democracy                     | 280 |
| 6.   | The Institutions of Experimental Democracy         | 285 |
| 6.1  | Introduction                                       | 285 |
| 6.2  | The Identification of Problems                     | 290 |
| 6.3  | Creating Diversity in Problem-Solving              | 297 |
| 6.4  | Making Feedback Effective                          | 330 |
| 6.5  | Putting it All Together                            | 345 |
| <br><i>Bibliography</i>                                |  | 349 |

## List of Figures and Tables

|  |     |
|--|-----|
| Figure 1: Democratic Differentiation   | 15  |
| Table 1: Case 3                        | 125 |
| Table 2: Case 3a                       | 126 |
| Figure 2: A Complex Estimation Problem | 251 |

## Acknowledgements

This dissertation would not have been possible without the help and support I got from a great number of people. First I would like to thank my advisors at Columbia University, Jon Elster, Melissa Schwartzberg, and Philip Kitcher, who have been unfailing in their support throughout this project. Jon Elster stoked my interest both in the problem of uncertainty in politics, and in the questions of the design of collective decision-making institutions. His guidance and inexhaustible knowledge have been very important for me. To Melissa Schwartzberg I owe great debts of gratitude for her close mentorship. She followed my project at every stage, and without her practical advice as well as constructive criticism this would not have been successful. Philip Kitcher introduced me to an exciting great body of research at a crucial stage of my project, and I have learned (and hope to learn more) from his kind and perceptive comments. I also thank David Johnston and Jonathan Bendor for agreeing to be on my dissertation committee and for their comments and criticisms, which have helped me enormously (and will do so in the future) to clarify my thoughts on this project.

I am very grateful to the rest of the faculty and especially also the staff at the Department of Political Science at Columbia, for their uncomplicated but crucial support. I would also like to thank Jeffrey Lax and Yotam Margalit for agreeing to be on my proposal committee and for their comments, which set me on the right track in the early stages of the project. I also owe thanks to the Mellon Interdisciplinary Graduate Fellows program at Columbia's ISERP (now INCITE) for their support, material and—especially—otherwise.



At Columbia and Oxford, throughout the years there, I benefited from the help, support, and criticism of teachers, friends and colleagues. Among many more, they are: Hiroaki Abe, Ashraf Ahmed, Andreas Avgousti, Adam Bronson, James Chappel, the late Jerry Cohen, Simon Collard-Wexler, Axel Domeyer, Kevin Elliott, Aimée Genell, Rob Goodman, David Hine, Kate Jackson, Jennifer Hudson, Yao Lin, Luke MacInnis, William McAllister, Alex Rigas, Pierce O Reilly, Neelanjan Sircar, Peter van der Windt, Jonathan Wright, Liya Yu and Jim Zetzel, as well as those I have likely forgotten to list here. Special thanks also to Michelle Chun for reading and commenting on almost the entire manuscript.

Parts of this dissertation were presented at the following conferences and workshops (in chronological order): Brave New World 2011 in Manchester, the Association for Political Theory Annual Conference 2011, the Oxford Political Theory Graduate Conference 2012, the Midwestern Political Theory Annual Meetings 2012 and 2013, and the American Political Science Association Annual Meeting 2013. I owe thanks to all discussants, panelists, and audience members who were so gracious as to give me comments and encouragement.

My family, both in Germany and in the United States, has been a constant source of support in so many ways, and I owe them more than I could express here.

Most of all, I thank Solongo Wandan. None of this would have been remotely possible without you being part of my life. Thank you.

*For Solongo*

## Introduction: What Should a Theory of Democracy Do?

### 1. An Icelandic Introduction

Let me begin with a political story that, while it concerns a small and—to many people—probably peripheral place, illustrates a number of key themes of this dissertation.<sup>1</sup> In the wake of the great banking crisis of 2008 the government of Iceland faced a particularly difficult decision. Icelandic banks had actively sought investments from foreign individuals, especially other Europeans in order to finance their investment in the immensely risky derivatives business whose collapse would eventually cause the crisis. The “Icesave” interest-bearing savings accounts from the Icelandic *Landsbanki* became a symbol of this business model. They were advertised heavily abroad and promised high returns on savings. Subsequently, as we all know, the Icelandic banks did fail, and the people who had invested in these accounts lost their savings.

---

<sup>1</sup> This story was reported on “Planet Money” on National Public Radio, (NPR) in 2011. All the quotations in this section are transcribed by the author from the radio broadcast: David Kestenbaum and Baldur Hedinsson, “A New Mom, Bjork’s Dad and the President of Iceland” (NPR: Planet Money Podcast, National Public Radio, 15 Apr. 2011).

Most of the investors, who came mainly from the UK and the Netherlands, were “bailed out” by their own respective governments (i.e. their taxpayers covered their losses). The question facing the government of Iceland was this: should an effort be made to reimburse the governments of the UK and the Netherlands for the bailout of their citizens who sustained losses in the collapse of the Icelandic banks, or is that not Iceland’s responsibility? This is a question with pragmatic as well as moral dimensions. On the one hand, the credibility, respect and standing of Iceland within the global economic system were at stake—which can have extreme consequences for such a small country that relies on international trade. Furthermore, perhaps this was also the “right” thing to do in a moral sense, given the massive financial loss for citizens abroad.

Eventually, a deal was reached between Iceland, the UK and the Netherlands. This deal was intended to guarantee Iceland’s standing within the global economic community, and to some extent to take responsibility for the irresponsible investments made by the Icelandic banks. The total amount of the deal was relatively small, but due to the tiny size of the Icelandic population, the amount to be paid to foreign governments by the Icelandic citizens was equivalent to roughly US\$60,000 per capita; a very substantial sum, especially for a country in the middle of an utter collapse of its entire financial sector.

Eventually the Icelandic parliament accepted the deal. Evidently, it weighed the concerns for Iceland’s standing with respect to the other countries and the concern for the sovereign debt rating of Iceland—and perhaps also the moral responsibility for foreign citizens—more heavily than the immediate cost to the taxpayer.

However, the (figurehead) president of Iceland at the time, Ólafur Ragnar Grímsson, used one of the office's residual privileges (*a de jure*, but not really *de facto* right) and vetoed the parliament's decision, thus throwing the decision to a nationwide referendum. No Icelandic president before him had ever used this power—or indeed exercised any legislative influence whatsoever. The president felt it necessary that the population should make this decision. After only a very short run-up, the referendum reversed the decision by the parliament and the deal fell through. The majority of the people of Iceland did not want to reimburse foreign governments. Evidently, they weighed the moral, legal, and economic concerns quite differently than the parliamentarians.

Of course, a debate arose immediately whether the decision to throw this issue to a popular referendum was a good idea. Illustrating the complexity of the problem, the reporters chronicled the sincere attempts of a citizen of Iceland, , in the run-up to the referendum, to make heads or tails of the issue. She consulted with economists, international lawyers, and fellow citizens, often receiving contradictory answers; often being confounded by the complexities of the issue, and perplexed by the amount of time citizens would need to expend on *research* to reach the point of being able to make an informed decision on this issue:

“A lot of Icelandic people are well-educated and feel obliged to know what they are voting on. I mean, this is an important decision ... But people are maybe doing it in whatever spare time they have ... the president of course asked us to do this, but he is not paying us to do it ... I think in many ways it would have been better to get somebody who is professional.”<sup>2</sup>

---

<sup>2</sup> *ibid.*

Her concern was not primarily that people are not *capable* of gaining an understanding of the complex problem, but that they should be able to *delegate* that task to people who do that professionally. One may also wonder whether most people in Iceland or indeed anywhere could be expected to be anything as diligent in informing themselves about such weighty issues as she was.

On the other hand, President Grímsson's attitude exhibited a degree of mistrust for representative institutions in general, and of their capability to get this specific decision right in particular. He expressed his doubts that parliamentarians would do any better than the people at large in gaining a clear understanding of the issue; more generally he remarked in an interview:

“You have now experienced how you come to a conclusion in a complicated issue. So I think, and I definitely hope, that in the future you will never let anybody tell you that this issue or that issue is so complicated that you should simply trust them to take a decision. Because democracy, fortunately, is a system where the farmer and the fisherman has the same right as the president ... First of all, on most issues we let those who serve in the parliament, or the cabinet, or the city council take the decision. You say, and you listed all these complicated things and so on, and the normal person can't deal with it. But: are you sure that the members of parliament can do this as well? The most important lesson I have learned throughout my public life is that people are not stupid; and you can rely on their wisdom in the long run, more than on the so-called experts.”<sup>3</sup>

The president's motivation was at least partly the idea that the decision made by the citizens would be *better* for them and for Iceland, and perhaps better in an objective sense, than the decision that was made by the Icelandic parliament.

This story, in the context of the Icelandic microcosm, contains many of the issues and illuminates many of the faultlines and problems of contemporary democratic politics. First, there is the question whether important political decisions should be taken through

---

<sup>3</sup> *ibid.*

referenda (or other more direct ways of democratic influence) or should “professional” elected politicians make them? Should there be a greater role for direct participation even with respect to questions as complicated as the reimbursement issue, or should there be more democratically authorized delegation to people who professionally deal with them? How would we even go about answering these questions? These and many similar questions about how democratic institutions should be organized are explicitly and implicitly behind many political disputes today. Political theory, and democratic theory in particular—as I understand it—should be able to offer answers to these questions. Of course both decision mechanisms in the Icelandic case can plausibly be called *democratic*, and therefore are *prima facie* democratically legitimate. Therefore the argument that one of them is inherently more democratic than the other does not seem to be available; there must be some more substantive normative standard. So the question is *which* normative criteria we could find to *differentiate* between democratic systems and individual democratic institutions.

Second, the president’s response highlights that the *quality* of decision-making, or the capacity to address highly complex questions is an important factor when we think about political decision-making. It seems that for many people, what we think about democratic processes does not depend on the nature of democratic “input” procedures alone. For many, having an equal right to vote—whether in elections or referenda—is not by itself enough for a system to be considered desirable: if we cannot also trust in the system’s capacity to deal with the large and small problems that arise in the course of social life, then we may have no reason to trust in its legitimacy. The “output” of democratic decision procedures should have some role in this evaluation.

Finally, the issue of the Icelandic reimbursement deal illustrates another key feature of political problems today: *complexity*. Because this issue has so many dimensions and unknown variables, it is *difficult* to clearly figure out what should be done. Crucially, the example is also an issue where *experts* disagree among themselves just as much as “ordinary” citizens. Predicting what would happen either way depends on a host of factors that are themselves unpredictable: the reaction of the UK and Dutch governments (and citizens), the reaction of the “financial markets” (themselves made up of a multitude of independent but adaptive actors), the status of international law and how it will be interpreted by courts and scholars, and last but not least, what the *moral* consequences of the action will be. The ongoing global financial crisis and recession since 2008 is not the only area that is complex in this sense. Many policy fields, global or local, are riddled with profound *uncertainty*. The problem is not that we basically know what to do and that the problem is how to get the political system to get there. A more fundamental problem is that we (often) do not know what to do.

Putting these notions together, we are of course faced with somewhat of a quandary. We want to know whether President Grímsson’s decision was justified, and for that, it seems, we have to know whether accepting or rejecting the deal was the right thing to do. But because the reimbursement question is so difficult, we do not know what the right decision would be (right for Iceland, or right in an even more objective sense). As noted above, experts were entirely divided about this, as were the citizens. And therefore we do not know whether the decision was justified; and so on.

This is—in basic terms—the topic of this project. Its stated goal would be to define a theory of legitimate democratic authority that is sensitive to people’s concerns



with the quality of decision-making, *under conditions of pervasive uncertainty*. In practical terms, this means coming up with a *standard* or a *criterion* of legitimacy to evaluate different forms of democratic decision-making. This standard should include at its core the *capacity to adequately resolve complex political problems under conditions of uncertainty*. Such a theory therefore can tell us not only whether democracy is better than other political systems (though it does that as well), it also tells us *which concrete forms* of democratic government are actually advantageous—more specifically, which functions a good democratic system should fulfil. It gives us a critical viewpoint from which to evaluate our own democratic system.

It seems to me that this addresses some central concerns of political life in democratic countries today. We can frequently see these kinds of conflicts of principle between parliamentary/representative, administrative or “expert” decision-making (what is often called “technocratic”) on the one hand and more participatory alternatives, on the other hand. At the European Union level, for instance, this conflict became apparent with the French and Dutch “No” to the constitutional treaty in 2005, after which the governments of the EU pressed on and passed very similar constitutional changes under the different name of “Lisbon Treaty.” Thus, as some people think, ignoring the verdict of the people. The ratification of the Lisbon Treaty led to another conflictual episode when the Irish people rejected it by referendum, only to have the referendum repeated a second time (presumably because this was the “wrong” outcome).

In Colorado and Washington, 2013 saw the legalization of marijuana through ballot initiatives, in clear contravention of the majority wishes of the respective state legislatures—let alone the majority opinion of Congress at the Federal level. Furthermore, in a number of American states the debate about same-sex marriage has devolved into a

back-and-forth between the outcomes of referenda, legislative decisions and judicial verdicts at various levels. California's Proposition 8 is of course the key example of this.

One might think that for better or worse, this "crowdsourcing" of political decision-making and the bypassing of the traditional representative democratic fora has become somewhat of a trend. To be sure, often these referenda are used merely to advance a particular agenda, not for the objective reason that they might make better decisions than the tried-and-true institutions of democratic politics, but sometimes, as we have seen with President Grímsson, the reason given for crowdsourcing is the advantage the processes have with respect to the quality of their decisions.

This fits well within a wider trend. Riding the same wave, in academic and pop-scientific literature about social epistemology the "wisdom of crowds," which is the title of James Surowiecki's bestseller, has become somewhat of a buzzword.<sup>4</sup> "Crowdsourcing" all sorts of epistemic tasks, especially through new technology, has become omnipresent, and indeed shows some striking successes. From a crowd estimating the weight of an ox at a state fair, which is the frequently cited example of early statistician Francis Galton, to analyzing data, utilizing the collective intelligence of everyone combined, so it seems, often outperforms what any one individual or selection of individuals could achieve.

But what should we think about *political* decisions? For explicitly political, complex decisions, such as the long-term consequences of the Icelandic reimbursement deal, what can we expect the wisdom of crowds to achieve? How wise is it when we apply it to a broad decision such as this? And more importantly, can we systematically understand

---

<sup>4</sup>James Surowiecki, *The Wisdom of Crowds*, 1st ed. (New York: Doubleday, 2004).

when a decision such as President Grímsson’s is adequate and when it isn’t? And given these limitations, can the wisdom of crowds actually justify democracy and democratic decision-making? And just as importantly, *which* kind of democratic system could actually be justified? As already mentioned, both the yes and the no decision to the reimbursement deal could plausibly be called “democratic”—a representative mechanism decided one way, and people in a referendum another.

## 2. Differentiating Democracy

This project is motivated by the general issues behind the Icelandic example: many of us would like to improve the quality of political decisions in our democratic systems, but there is so much uncertainty and disagreement about what this “quality” is that it is radically unclear how we should do so. In other words, while we care very much about the *outcomes* of the democratic process we don’t know *ex ante* how to get them (and sometimes not even what they look like). What one citizen sees as a just political goal may be a grave injustice for another, and for political goals that command widespread agreement, there is often deep uncertainty about which policies will actually turn out to *work*.

Consequently, if we do not quite know what “better decision-making” is, it is unclear which institutional reforms or changes in democratic structures would actually promote it. Should there be more direct democracy, or less? Should new entrants into the political sphere—like new parties—receive additional support, or should they be treated on equal terms with the established players? Democracy is a particularly wide umbrella

term, and clearly not all institutional constellations and rules and regulations that can be called democratic function equally well.

Therefore, in this section I offer a few words about how this study fits into the landscape of democratic theories. Now, one might wonder whether such apparently practical questions of how to design institutions really have a place in the fundamental debate about which forms of political system are normatively desirable. One might think, as many political philosophers do, that the quality of political outcomes should not really enter into the justification of political authority (even though some think otherwise). However, in keeping with the rise in popularity of the idea of the “wisdom of crowds” in other fields of inquiry, in recent years we have seen a revival of “instrumental” or “epistemic” theories of democracy. These share the fundamental principle that political systems should be considered legitimate if and only if they tend to produce good political outcomes (and illegitimate if they systematically fail to do so). This is coupled with an argument that democracy has a specific edge over its alternatives when it comes to producing these good outcomes. Examples of such arguments abound.<sup>5</sup>

---

<sup>5</sup> see for instance David M Estlund, *Democratic Authority: a Philosophical Framework* (Princeton, NJ: Princeton University Press, 2008); Hélène Landemore, *Democratic Reason* (Princeton, NJ: Princeton University Press, 2012); Scott E Page, *The Difference* (Princeton: Princeton University Press, 2007); Josiah Ober, *Democracy and Knowledge: Innovation and Learning in Classical Athens* (Princeton, NJ: Princeton University Press, 2008); Josiah Ober, “Democracy’s Wisdom: An Aristotelian Middle Way for Collective Judgment,” *American Political Science Review* 107.1 (2013): 104-122; Elizabeth Anderson, “The Epistemology of Democracy,” *Episteme* 3.1-2 (2006): 8–22; Elizabeth Anderson, “Democracy: Instrumental vs. Non-Instrumental Value,” in *Contemporary Debates in Political Philosophy*, ed. Thomas Christiano and John Christman (Oxford: Wiley-Blackwell, 2009), 213–227; Christian List and Robert E Goodin, “Epistemic Democracy: Generalizing the Condorcet Jury Theorem,” *The Journal of Political Philosophy* 9.3 (2001): 277–306; Robert E Goodin and Kai Spiekermann, “Epistemic Aspects of Representative Government,” *European Political Science Review* 4.03 (2011): 303–325; Franz Dietrich and Kai Spiekermann, “Epistemic Democracy with Defensible Premises,” *Economics and Philosophy* 29.01 (2013): 87–120.

Normative justifications of political systems more generally must have a two-stage structure.<sup>6</sup> If we want to justify a particular political procedure A, we have to firstly abstractly determine which features we consider normatively desirable in a political system, and then secondly argue that system A (uniquely) exhibits these criteria.<sup>7</sup> We have to offer some *grounds* for the legitimate authority we are claiming for A, and then give reasons how A respects or fulfils these grounds. Justifications can therefore break down at two points: when the grounds we propose do not turn out to be sufficient to justify coercive political authority, *or* when the secondary claim that our preferred system is justified on those grounds, fails. Thus, for example, a defense of particular democratic institutions that is premised on the normative value of political equality can break down if it is shown that political equality is not a sufficient value to ground authority, *or* if it is shown that the preferred form of democracy does not actually realize political equality in the way intended.

Justifications of democratic authority, especially also epistemic justifications, often tend to be framed as a *defense* of democracy “in general” against its alternatives: against monarchy, authoritarianism, Leninist-style communism, or anarchism; in other words, they are attempts to establish the legitimate authority of democratic systems against explicitly non-democratic alternatives. Even though this is of course an important task, this leaves open a lot of questions.

---

<sup>6</sup> Richard J Arneson, “The Supposed Right to a Democratic Say,” in *Contemporary Debates in Political Philosophy*, ed. Thomas Christiano and John Christman (Oxford: Wiley-Blackwell, 2009), 197–212.

<sup>7</sup> A justification of some political system needs to show this uniqueness if it is supposed to be an overriding justification. For a (mere) *pro tanto* justification it may be enough to show that the criteria are sufficiently satisfied.

As I argue, if we examine the nature of the *grounds* given for political legitimacy, it will become clear that they do not justify democracy in all its forms, but only justify those democratic systems for which a convincing case can be made that they are actually effective (in relative terms) at bringing about the desirable outcomes.

The grounds commonly cited in support of epistemic justifications of democracy may not support all and every institutional incarnation of democracy. A commitment to an instrumental-epistemic theory of political legitimacy forces one to open the box “democracy,” and look which one of the items inside, or which combination thereof, is actually justified as a political system. My answer, as given in this dissertation, is that under those conditions, what we should seek is a form of *experimental democracy*. Further below I will briefly outline what I mean by this term.

The basic thought is that there is no single essential form of “democracy,” and many real existing democracies are not democratic “all the way down,” i.e. not all of their institutions or decision-making procedures are actually themselves democratic. Most democracies actually contain a *variety* of first-order decision-making agents, overtly democratic ones and those less so. All or almost all constitutional democracies around the globe already employ a mixture of those. The most obvious point is that despite being labeled “democracies,” actual political decisions are usually not made by the people at large,<sup>8</sup> but through a variety of representative institutions, judicial review, bureaucratic agencies, public-private partnerships and other elements that tend to be present in such systems.

---

<sup>8</sup> With exceptions, of course.

It seems that that many epistemic defenses of democracy, as they stand, tend to give little normative guidance as to which of these combinations is in fact justified. However, theories of democratic legitimacy—including instrumental ones—should provide a critical viewpoint that not only justifies democracy against non-democratic alternatives, but also tells us which mechanisms of democracy we should endorse. They should give us an answer to President Grímsson’s problem. To put it negatively: there are political systems that are pretty bad from an epistemic viewpoint but are nevertheless democratic—I think outcome-focused theories of democracy should have something (critical) to say about that; and indeed I think that they do.

One reason is that the focus on defending democracy “in general” has directed attention to the dichotomy between democracy and some clearly inferior non-democratic alternatives: like the rule of philosopher kings or otherwise self-proclaimed experts. The arguments conclude on establishing that that democracy is shown to be superior, on epistemic grounds, to any of these other forms of political organization. The main concern of some instrumental-epistemic justifications of democracy in particular sometimes seems to be a reconciliation between our pre-theoretical normative commitment to democracy and our intuitions about the perceived lack of competence of the general public,<sup>9</sup> rather than the creation of a critical standard to see the value of democracy as such. All the while the equally interesting question of which concrete form of democracy in particular may be justified by epistemic arguments, and may be superior to non-democratic alternatives receives too little attention. In particular, some forms of

---

<sup>9</sup> Michael Fuerstein puts this point succinctly, in form of a question: “how can democratic governments be relied upon to achieve adequate political knowledge when they turn over their authority to those of no epistemic distinction whatsoever?” See Michael Fuerstein, “Epistemic Democracy and the Social Character of Knowledge,” *Episteme* 5.01 (2008): 74.

democracy might turn out to be rather bad—indeed possibly worse (on some counts) than the dreaded epistocracies.

Let me briefly emphasize this point on the basis of David Estlund’s theory. Estlund has probably been more influential than any other contemporary political philosopher in reviving epistemic theories of democratic authority, and the theory will be discussed in much more detail in chapter 2.<sup>10</sup> Without going too deep into this issue here, his theory on political legitimacy, *epistemic proceduralism*, argues that political legitimacy depends on a regime’s reliable tendency to produce good outcomes, provided that the claim of reliability is acceptable from all qualified viewpoints. Roughly speaking, the criterion for legitimacy is whether a system produces more correct decisions than a random decision procedure (subject to qualified acceptability). Now if we want to justify democratic systems “in general,” we have to establish another premise: that all systems that can be characterized as democratic will be better than random, and indeed that there is no meaningful difference between such different democratic systems. However, just as we can easily imagine democratic systems that are better than random, we can easily imagine theoretical cases of terrible systems that nevertheless are democratic.

There are of course reasons why we might think that democratic mechanisms might lead to epistemically beneficial outcomes, but no reason why some variants of democracy might not be worse even than random. The situation, for Estlund, then, might well look like in Figure 1 below. Here the range of democracies contains structures both that are better and that are worse than random procedures. If this is the case, only those

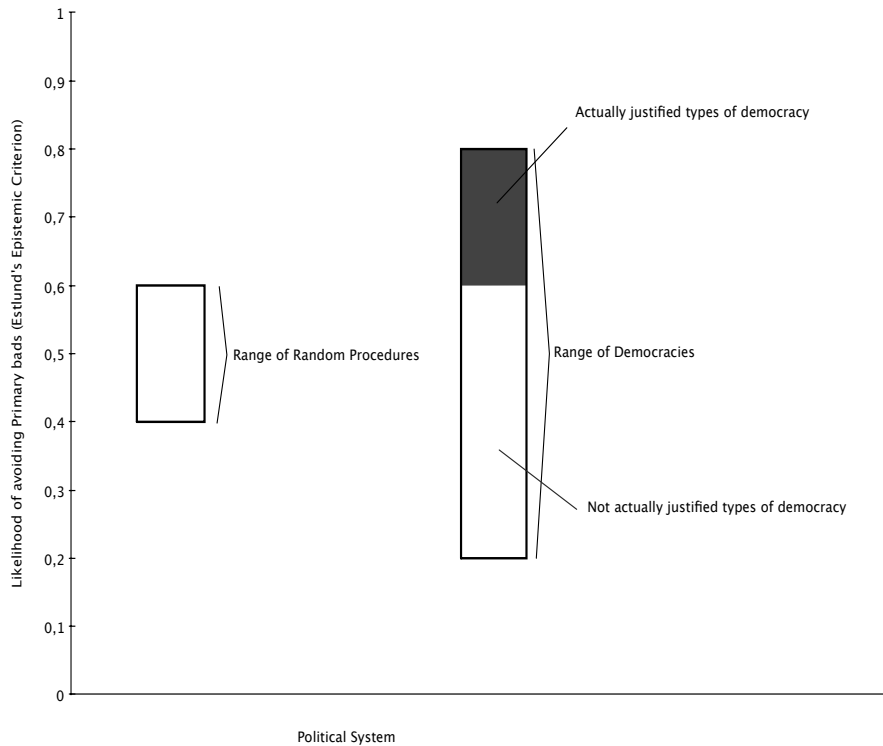
---

<sup>10</sup> Estlund, *Democratic Authority: a Philosophical Framework*.



types of democracy that are in fact better than random will be justified. The dark grey section on the right in Figure 1 represents this range. All other forms of democracy, despite being democratic, are not actually desirable.

But this is exactly what we need: a criterion that normatively differentiates democratic institutional forms, rather than a wholesale “defense” of all systems labeled “democracy.” The question whether democracy is better than non-democratic options such as selection of rulers by random sortition, let alone tyrannical or authoritarian alternatives, is no more important, from a justificatory point of view, than the question which concrete democratic institutional setup is good enough to be instrumentally justified.



*Figure 1: Democratic Differentiation*

I would even say that at least in the context of the current Western industrial democracies, the latter “practical” question is in any case arguably more “alive” than the former: beyond the examples already mentioned large conflicts persist between advocates of referendum decision making and advocates of exclusive parliamentary decision-making; between proponents of unelected elements within democratic systems (like the Federal Reserve, the U.S. Supreme Court, the House of Lords, the UK Equal Pay Commission, etc.) and advocates of a restoration of parliamentary superiority over all those policy fields;<sup>11</sup> advocates of strengthening executive government vis-à-vis other branches of government,<sup>12</sup> or supporters of the European Commission’s “benign elitism”<sup>13</sup> versus advocates of strengthening parliamentary oversight and control; advocates of tighter restrictions on voting rights<sup>14</sup> versus advocates of expanding the franchise to the under-18s or alien residents.<sup>15</sup> What is remarkable is that a preference for “non-democratic” elements tends to be coupled with an acceptance of “democracy” as a general ideal.<sup>16</sup> Epistemic theories of political legitimacy, as I argue, not

---

<sup>11</sup> Colin Crouch, *Post-Democracy* (Cambridge: Wiley, 2004).

<sup>12</sup> Eric A Posner and Adrian Vermeule, *The Executive Unbound: After the Madisonian Republic* (Oxford: Oxford University Press, 2011).

<sup>13</sup> The Austrian novelist Robert Menasse has recently launched, across several media outlets, a much-debated defence of the European Commission, arguably the most elitist and least democratic (in a populist sense) legislative agent in European Union policymaking. He asks whether it “is not rather the case, at this point, that we should admit that it is progress, and indeed an emancipation, when the basic conditions of our life are no longer decided upon by popular elections.” Robert Menasse, “Joseph und Angela: Menasse über die EU,” *Kleine Zeitung*, March 27, 2010. My own translation.

<sup>14</sup> for instance Jason Brennan, *The Ethics of Voting* (Princeton, N.J.: Princeton University Press, 2011); Jason Brennan, “The Right to a Competent Electorate,” *The Philosophical Quarterly* (2011).

<sup>15</sup> Andrew Rehfeld, “The Child as Democratic Citizen,” *The Annals of the American Academy of Political and Social Science* 633.1 (2011): 141–166; Douglas A Chalmers, *Reforming Democracies* (New York: Columbia University Press, 2013).

<sup>16</sup> For instance, there seems to be a desire among Americans to be ruled by competent, non-self-interested people, rather than for self-rule via increased participation. See John R Hibbing and Elizabeth Theiss-Morse, *Stealth Democracy: Americans' Beliefs About How Government Should Work* (Cambridge: Cambridge University Press, 2002).

only enable us to have answers to these questions, indeed they commit us to find answers to these questions. Epistemic theories of political legitimacy are critical tools, and epistemic theories of democracy are theories that are, by nature, critical of democracy as well.

### 3. The Argument and the Level of Analysis

In this dissertation I offer a model of democracy—“Experimental Democracy”—that accepts the epistemic view that the quality of decisions matter, but that defines this quality by taking into account the circumstances of uncertainty and disagreement. On this account, a desirable political mechanism is one that realizes an *experimental method* of policy-making directed at solving problems, such that we can expect it to make progress over time, even though we cannot rule out that it will get things wrong—possibly even frequently. I also show how democracy is most likely to realize such an experimental method (compared to feasible alternatives), and which particular institutional features may serve this purpose.

The dissertation proceeds, as it were, from the bottom up. Before we arrive at the full model of Experimental Democracy, we have to start small and figure out why an experimental method of policy-making is the appropriate response to uncertainty and disagreement in general, and whether and how it is really normatively desirable. We have to come up with an instrumental theory of political justification precisely for conditions of uncertainty. This is the subject of Part I of the dissertation. Part II then turns to *democratic* theory in particular and more clearly defines Experimental Democracy in order to give us the critical standard that can be used to differentiate democratic types. This part then

turns to the institutional details, giving us some indication of how to answer the general questions mentioned above.

A note on methodology and assumptions: throughout I aim to follow Rousseau's principle to take "men as they are and laws as they might be." In order to figure out the normative value of particular concrete institutional configurations, I use *standardizing* assumptions about human behavior, but try to refrain from *idealizing* assumptions about either the cognitive capacities or moral motivations of people. The typical democratic agent in my picture is basically rational, but not a selfless martyr, she has not terribly much time to think about politics and limited access to information. She is, on the other hand, neither completely irrational nor willfully out to harm other people for no gain of her own. One might, therefore, categorize this as non-ideal theory, depending on one's picture of what people are "really" like.

Another note, however, on the limits of the dissertation. The goal of this dissertation is not to come up with concrete recommendations or an institutional blueprint for experimental democracy. I do not attempt to give very detailed guidance on specific institutions and rules that should govern collective decision-making. Instead the goal is, on a somewhat more abstract level, to come up with *principles that ought to govern such choices*, that is to define which principles for institutional design an experimental system of government would embody (and why that would be desirable).

The reason for the limited goal of this dissertation is that the real effect of concrete decision structure designs on policy choices probably depends on many contingent or context-specific circumstances and is subject to case-dependent practical constraints.

Therefore it is a subject of empirical investigation that lies outside the scope of this dissertation. Given reliable empirical findings about these questions, one should be in principle able to apply the principles of experimental democracy—outlined in Part II—to concrete institutional reform.

Consider an example: the proceedings within a decision environment can be transparent or closed to the public. The effect of this choice, however, depends on the contingent public opinion structure, as well as the dependency relations between a representative and his constituency, in a given society. Take the case of a transparent legislative process. When a representative's constituency is wide-spread and diverse, she may be compelled to frame her arguments in the public interest—and we may consider that a good thing. If however, her constituency values the appearance of “being tough,” and of fighting for their narrow interests, publicity may drive her to adopt an obstructionist position that leads to bargaining or to inefficient policy-making. The effect of a design recommendation (such as “guarantee a high degree of publicity”) therefore seems to depend on empirical circumstances specific to a given society.

However, independently of this question about the effect of transparency we can debate whether the rationale given for transparency (the principle of institutional design which is supposed to be realized by transparency) is actually a *good* one. We might wonder whether the *principle* governing institutional design of collective decision structures should be to maximize the occurrence of public-spirited argument or bargaining. This is the level of analysis at which experimental democracy is pitched as well. In part II of the dissertation I provide *principles of how to organize decision-making*, not concrete recommendations for institutional rules that realize these principles in particular

situations. I will give a number of concrete examples of institutional mechanisms that could be desirable on the basis of experimental democracy, but I will not stake a claim that they will always and under all circumstances do so.

#### 4. The Normative Basis

Before moving on to an outlook of what lies ahead, let me just say a few words on the most fundamental normative assumptions underlying this project. I construe the argument as a question of *justifying political authority*. This of course assumes that political authority needs justification at all; but beyond that I assume that any political authority has to be justified *to the people* living under it, if not explicitly, then in a way that they could reasonably accept. I am therefore firmly in the realm of *public reason liberalism* in the Rawlsian formula,<sup>17</sup> or as Gerald Gaus calls it, *justificatory liberalism*.<sup>18</sup>

The problem of political justification arises from the fact that political systems are imposed on individuals, usually without their consent. After the imposition, the system (or the individuals authorized by it) claims *authority* over citizens' lives. In this view, political authority is legitimate if and only if this imposition can be *justified* to the people subject to it. One popular way to justify such an imposition has historically been to argue that somehow it is not *really* an imposition, and that citizens would actually consent if they

---

<sup>17</sup> John Rawls, *Political Liberalism* (New York: Columbia University Press, 2005).

<sup>18</sup> Gerald Gaus, *Justificatory Liberalism : An Essay on Epistemology and Political Theory* (Oxford: Oxford University Press, 1996); Gerald Gaus, *The Order of Public Reason : a Theory of Freedom and Morality in a Diverse and Bounded World* (New York: Cambridge University Press, 2011); see also Steven Wall, "On Justificatory Liberalism," *Politics, Philosophy and Economics* 9.2 (2010): 123–149; Steven Wall, "Public Reason and Moral Authoritarianism," *The Philosophical Quarterly* 63.250 (2012): 160–169.

could—or that in some sense, their behavior can actually be taken as consent. This is the tradition of the social contract, actual or hypothetical, that reaches back at least to Hobbes and Locke. Much theory has been written about how to interpret citizens’ acquiescence as consent. But at least since Rawls, the problem has more often been interpreted as one of coming up with a standard of reasonableness such that absent countervailing ideas, reasonable citizens *would* hypothetically agree to such a contract. The basic Rawlsian argument is that reasonable citizens can be considered as agreeing as long as the hypothetical contract is *fair* in some sense. This of course leads to the next questions of what that means. As Charles Beitz puts it, “fair terms of participation are those that no citizen has a sufficient reason to refuse to accept, given that everyone shares a desire to come to some agreement on some mechanism for participation. The main difficulty in working out this idea is to explain what could count as a sufficient reason for refusal.”<sup>19</sup>

This general form of argument assumes a default condition of people as free to do what they please, without legitimate interference. As Gaus puts it: “Freedom to live one’s own life as one chooses is the benchmark or presumption; departures from that condition—where you demand that another live her life according to your judgments—require additional justification.”<sup>20</sup>

Such a justification must be based on a *reason*: “A reason R is a moral, impartial, reason justifying x only if all fully rational moral agents coerced by x-ing would

---

<sup>19</sup> Charles R Beitz, *Political Equality* (Princeton, N.J.: Princeton University Press, 1989) p. xiii.

<sup>20</sup> Gaus, *Justificatory Liberalism : An Essay on Epistemology and Political Theory* 165.

acknowledge R, when presented with it, as a justification for x-ing.”<sup>21</sup> I accept this mode of argument. I also do not have the space here to defend this basic assumption; let me just mention that much of normative democratic theory nowadays seems to argue from this basic starting point, as do I. My argument differs from Rawls and Beitz, though, as to what I argue can count as a *reason* such as may be given to a reasonable citizen, and which she would or could accept.

In particular, I accept *instrumental* reasons as a possible basis for legitimate political authority. In this way, we might think the best way to justify a political system to its subjects would be to refer to the expected quality of its *outcomes*, that is the anticipated consequences (over time) of its authoritative decisions. In particular, we might think that a system can be justified if it tends to make the *right* decisions: this is the idea of an epistemic justification of democracy. There are, of course, potential problems with such an account. But then the dissertation still has some way to go after this Introduction.

## 5. The Argument Coming Up

As already mentioned, the dissertation is divided into two general parts, which contain three and two chapters, respectively. Part I develops a theory of political legitimacy in the sense outlined above: a theory that is sensitive to the *outcome* concerns that many people share, but recognizes the fundamental disagreement surrounding this question. This is what I call a *pragmatic* or *experimental* theory of political authority. Part II then develops the theory of experimental *democracy*, and shows more clearly which form of democratic

---

<sup>21</sup> Gerald F Gaus, “Liberal Neutrality: a Compelling and Radical Principle,” *Critical Review* (1994): 143.



decision-making is justified on the basis of this theory of authority, and which institutional mechanisms do the “work” of legitimizing it.

Chapter 2 establishes that outcome considerations *in some important sense* matter for our normative assessment of democratic mechanisms, while leaving open for now what that sense is. It abstractly establishes the importance of outcome considerations for the evaluation of different democratic procedures. This chapter proceeds in several steps. First, I argue against *pure proceduralism*, the view that political systems ought to be assessed only in terms of their intrinsic features, for instance the equal access to political power they offer. Second, I argue against what I call *lexical proceduralism*—a view I associate with David Estlund. This is the idea that *both* procedural and outcome considerations matter for our normative assessment of democracy, but that in cases of conflict, procedural considerations have lexical priority, or “trump” outcome considerations. The problem with this is that the reasons that accord priority to procedural considerations (mainly the fact of disagreement) at the same time undermine the secondary outcome considerations. Hence, it must be allowed that at least sometimes, outcomes may outweigh procedural values. Finally, I consider what I call the “default” argument for proceduralism. This view focuses on the *de facto* deep and pervasive disagreement about what constitutes a good outcome and argues that as a “default” we have to resort to procedural criteria like fairness or equality to assess democratic procedures.

In response to these views, I point out that the fact of deep disagreement does not imply that there can be no non-procedural criteria for what makes a better or a worse outcome. Nevertheless, even if outcome-based evaluations of democratic procedures are *in principle* possible, proceduralists are quite correct that disagreement and uncertainty

make it *de facto* somewhat problematic to come up with such a theory. The last section of this chapter clarifies this problem: we should see this disagreement as a source of *uncertainty*, and as such it can be grouped together with other possible sources of uncertainty. Since outcomes are important, democratic institutions ought to be measured by their capacity to make adequate decisions precisely *under such conditions of uncertainty*. The dilemma with which the chapter ends is therefore: how can we understand the outcome (epistemic) reliability of democratic institutions when there is uncertainty about the standard of good outcomes?

Chapter 3 provides the answer to this last question: under conditions of uncertainty, epistemic reliability should be interpreted not as the likelihood of a democratic procedure to “get it right” on as many political decisions as possible (to “optimize”), but as the *pragmatic* capacity to solve political problems as they come up, and perhaps in this way overcome important problems (over the long term), despite the fact that there may be failures along the way. This chapter begins by discussing how we should understand the idea of objectively *better* and *worse* outcomes even under situations of uncertainty, and/or unresolved conflict. How should we understand the “reliability” of democracy on such a basis?

First I consider the “standard” model of epistemic reliability, as it is used in most epistemic arguments for democracy. This model interprets *reliability* as capacity for “truth-tracking,” i.e. the average likelihood that the procedure will get any given political decision “right.” This way to look at it faces the insurmountable challenge that it presupposes that we know (or have some idea about) what a “right” decision is: but under the conditions of uncertainty identified in this and in the previous chapter we cannot

assume this. If justification should be premised on some substantive assertion of what is right, which could be reasonably rejected, it cannot work. Finally, I formulate an alternative view: on this pragmatic account, democratic procedures should not be evaluated by the likelihood that they get things right, but accordingly to the *methodology* employed in making decisions under uncertainty. The task of a political procedure, in this view, is not to maximize their quota of “right” decisions, or even to maximize the number of problems solved—because we do not reliably know ahead of time which procedure would do so—but to progressively resolve problems of uncertainty *as they come up*. On this model, democratic politics is more continuous with science or engineering, in general with our practical problem-solving activities, rather than with moral philosophy. The added benefit is that we already know from these fields that the best problem-solving strategy to deal with uncertainty is to *experiment*.

Hence, we should understand pragmatic reliability under conditions of uncertainty as *capacity to employ an experimental problem-solving strategy*. In this view we can evaluate different democratic procedures *pragmatically* according to *whether or not they fulfill the functional requirements of such an experimental strategy*.<sup>22</sup> This form of *pragmatic legitimacy* is my answer to the problem of how to think about the outcome-based quality of democratic procedures precisely when we do not know what the right outcomes are or how to get them.

Chapter 4 expands on the idea of experimental politics and the issue of if and how this may form the basis of a justification of political authority. The first part of this

---

<sup>22</sup> Subject to a feasibility constraint

chapter outlines in more detail what is required of an experimental strategy in politics. Here I draw on recent ideas of experimental methodology in other fields of inquiry. Basically, the key features of an experimental methodology are the creation of a variation of possible solutions, the controlled implementation, and especially a feedback mechanism to enable progressive adaptation. Next I discuss a possible objection: while it is clear how one may experiment with factual or “technical” questions (the “means” of politics), how can one experiment with normative/preferential questions—what we might call the *ends* of politics? I answer, with Dewey, that there is no bright line between “means” and “ends” when it comes to their evaluation. Far from considering them of ultimate value, we evaluate our ends in terms of the means required to attain them. The appropriate reaction to such a conflict of ends-in-view and necessary means, as Dewey points out, is *further inquiry* to resolve the conflict: either find different means to achieve the end-in-view, or discard it in favor of better ends.

The second question addressed in the chapter is this: how can we justify coercive experimental policy-making to the people living under the policy, given that it is at least problematic to coercively enforce an experiment on non-consenting persons? I suggest that when we consider the alternatives to this mode, “optimization” and a basically conservative-reactionary attitude, it becomes clear that they either systematically underestimate risks or overestimates the certainty with which we can foresee the consequences of our policies.

In summary, Part I therefore (a) identifies a key problem faced by democratic theory today: that we are interested in the quality of outcomes, but face uncertainty and disagreement about what this quality is; it (b) formulates a way to understand the

reliability of political procedures under those conditions: the pragmatic evaluative standard of democracy, which favors experimental methods of policy-making; finally, it (c) defends the experimental approach against some apparent ethical problems and shows how *democracy* in particular may be experimental.

Part II of the dissertation looks at *which* conception of democracy (and therefore which particular institutional manifestations of democracy) would be optimal according to the normative democratic theory developed in Part I. What does an ideal “Experimental Democracy” look like?

Chapter 5 discusses which general “model” of democracy would follow from the experimental conception. Experimental democracy requires (a) a separation of the governing function from the popular evaluation function (or the “creation of variation” function from the feedback function). On this basis I first discuss two models of democracy: democracy understood as judgment aggregation through majoritarian decision-making, and as a deliberative process interacting different viewpoints. I argue that neither of them is consistent with the experimental model, since they rely on the assumption that one first-order decision mechanism (namely a democratic one) is appropriate for *all* political decisions. That requires assumptions that we are not justified to make. Instead I argue for what I call the “control model” of democracy. In this model, we have to understand democratic activity as exercising *control* over the policy-making process (through setting incentives and sending information), *not* as directly exercising first-order decision-making. Therefore, political decision-making should be organized with its primary function in mind (the identification of and creation of possible solutions

to problems), and the public participation process should be organized with its own primary function in mind (feedback and control).

In the last part of chapter 5 I address the question of why, even if we accept all that I have said so far, we should prefer experimental *democracy* rather than some other experimental form of regime. The reason, as I argue is that democracy can be *robustly* expected to deliver the problem-solving we require: an experimental policy strategy requires accurate feedback on the effect policies have on actual people. While proxy measurements such as income per capita can of course be obtained non-democratically, a complete statistical picture of how policy affects people requires input from everyone affected. Secondly, while this information could also probably be gleaned through public opinion research, polls do not have the required causal effect on policy-makers: elections can (in principle) actually remove people from power, while people can hang on to power despite spectacularly low approval rates. Progressive problem-solving, as it were, depends on people's feedback to have a causal effect on the adaptive development of policy, not only on their ability to have their opinion counted.

Finally, chapter 6 considers what institutional features would be necessary to best approximate the functions of experimental democracy, both in terms of institutional requirements on the side of legislative activity, and in terms of the electoral side of politics. As mentioned, experimental democracy has three functional requirements: the identification of problems, the creation of a variation in solution proposals, and ex post evaluation through a feedback loop. I discuss in turn what these functions require in institutional terms. With respect to the identification function I argue in favor of lower barriers of entry for political contenders, as well as the limited use of submajority rules

within the decision-making body. The idea here is that we have to have mechanisms in place that reliably place problems (even those of minorities) onto the political agenda, and importantly, induce decision-makers to address them.

The variation function depends on the presence of wide cognitive diversity within the political decision-making process. Here I argue that not only is cognitive diversity compatible with the control model of democracy, they are actually mutually beneficial. Therefore, I have to argue against the view that employing cognitive diversity in decision-making necessarily implies first-order decision processes ought to be maximally democratic: in response to this I show that democratic mechanisms are not necessarily the optimal way to bring this diversity to bear on the policy process. If we want cognitive diversity, we have to *actively select* for this attribute in setting up our decision-making.

Finally, in terms of the feedback function there are two basic requirements: People have to be able to identify whether or not some policy is working or not, and they have to be able to effectively transmit this collective judgment to policy-makers. The first requirement calls for a basic norm of transparency. But since excessive transparency can obscure as much as illuminate, beyond this there should be strict *comprehensibility* requirements on the actions of the government. Ideally, policy should be short and easily understandable: but since in complex societies laws probably have to be complex, it would be a good idea to require the government to set out the *purpose* of each law in a generally accessible and comprehensible way, together with a *metric* by which one can decide whether the law is working or not. The second requirement faces the severe obstacle of the paradoxes of collective choice. Riker of course famously denies that we can

attribute any meaningful content to the outcome of a democratic vote at all.<sup>23</sup> A further problem with preference votes is that they carry too little meaningful information about how the public evaluates policies as such.

I argue, however, that these problems can be overcome. Here I focus on the method of *majority judgment* as it advocated by Michel Balinski and Rida Laraki.<sup>24</sup> This employs the judging method of assigning *grades* from areas such as wine-tasting and ice skating to democratic politics. This method is not subject to the important paradoxes of social choice, and removes (largely) the incentives to strategically misrepresent one's judgment. Furthermore, the collective judgment in this mode carries more useful *information*.

Experimental democracy, far from being an ideal conception only, therefore gives us a clear evaluative perspective on many practical institutional issues *within* the realm of democratic institutions more generally: campaign finance, methods of electing, transparency concerns, cognitive diversity, or the authority of experts or professional politicians vs. “the people.” As a whole, the dissertation attempts to reconcile the concern we have with the quality of political outcomes with the disagreement and extensive uncertainty we observe about precisely that, and to show how we should conceptualize democracy—and particular democratic institutions—on this basis.

---

<sup>23</sup> William H Riker, *Liberalism against Populism : A Confrontation Between the Theory of Democracy and the Theory of Social Choice* (Long Grove, IL: Waveland Press, 1982).

<sup>24</sup> Michel Balinski and Rida Laraki, *Majority Judgment : Measuring, Ranking, and Electing* (Cambridge, Mass.: MIT Press, 2010).



## Part I

---

### A Theory of Experimental Political Legitimacy

---

## The Value of Democracy

### 1. Introduction

The general problem I address in Part I of the dissertation can be summed up in terms of a dilemma: When trying to figure out whether and why democratic institutions are *better* than their alternatives (and which democratic institutions are “better” than others), one thing that seems to be essential is the *instrumental reliability*, i.e. the ability to produce *good* outcomes through its policy decisions, at least over the long haul and in general.<sup>25</sup> Contemporary politics is full of complex and difficult problems that have potentially enormous (beneficial or catastrophic) consequences. Political decisions have far-reaching and significant reverberations, and should not be taken lightly. However, at the same time it is clear that for many if not most political problems *ex ante* certainty about which decision will turn out to be right is not to be had. Assertions of “silver bullet” solutions to

---

<sup>25</sup> “Instrumental reliability” as I use it here is a measure (a standard) of a decision-procedure’s tendency to take *good* decisions, considered across the board and over the long run. As such, the definition is purely formal, as it leaves undefined what it means to “take good decisions.” This question is the subject of chapter 3. A similar standard is proposed by Arneson, “democracy is to be assessed by the consequences of its adoption and operation compared with alternatives ... one crucial standard for judging a society’s institutions and practices is the extent to which they are efficiently arranged to increase the likelihood that as time goes on our epistemic access to moral [and factual—*ed.*] truth will improve. See Richard J Arneson, “Democracy Is Not Intrinsically Just,” in *Justice and Democracy: Essays for Brian Barry*, ed. Keith Dowding, Robert E Goodin, and Carole Pateman (Cambridge: Cambridge University Press, 2004) 43.

political problems are, quite rightly, regarded as suspicious. But if there is substantial uncertainty and/or disagreement about what the right political goals are, and so no less about the appropriate means to get there, how can we develop dependable expectations about which political arrangements are likely to be reliable and which are not? How should we think about different variants of democratic decision-making in this context? How can we reconcile a demand for *reliability* of democracy with the fact of *uncertainty* about politics?

This whole first part of the dissertation looks closer at this problem. The present chapter asks *why* reliability should matter when we evaluate political procedures, and how that connects to the question of democratic political legitimacy. Many people agree that it is in some sense important that governments do the right thing. It is, however, not straightforward why that would be important, and what that would even precisely mean. Democratic theorists have produced a wide variety of arguments on what is valuable about democracy, and its reliability to produce the right outcomes is only one of them. And even within this school of thought, the idea of reliability itself can be interpreted in a wide variety of ways.

The following chapter 3 is then concerned with the question that immediately follows: how can we place such importance on “epistemic reliability” when we are unsure what the right thing to do even is? In the chapter I argue that indeed the only way to resolve the dilemma is by interpreting this idea *pragmatically*, or as I put it, according to the “scientist model” of epistemic reliability.<sup>26</sup> Understanding instrumental epistemic

---

<sup>26</sup> The term “scientist model” of politics already implies that I associate myself with a particular strand of pragmatist thought: in particular, the naturalist one that emphasizes the continuity of pragmatist thought

reliability as the capacity and tendency for *pragmatic problem-solving* is, as I show, appropriate for circumstances of pervasive and deep uncertainty that characterize politics today. Therefore, these first two chapters provide an alternative way to justify and evaluate different democratic regimes, one that is sensitive to outcome considerations while acknowledging the uncertainty and complexity we face in everyday political life.

But before we come to this argument, the present chapter should be seen as clearing the normative underbrush, as it were. I look at what it means to normatively evaluate different political decision structures and argue in favor of an *instrumental theory of democratic legitimacy*: the theory that the normative legitimacy of a political system is primarily determined by its capacity or tendency to produce the *right* outcomes, in some sense, through its decisions. Accordingly, in this view, *democracies are valuable in so far as and to the extent that they manage to produce good outcomes (generally, and over the long run)*; and they are not justifiable if they consistently fail to do so (at least compared to their alternatives). This is the very abstract sense in which I use “instrumental” here. Adopting Knight and Johnson’s term, we might also call this fundamental normative position “tempered consequentialism.”<sup>27</sup>

The main alternative view to this would be what we might generally call an *intrinsic* theory of democracy, which argues that democracy has value *qua* democracy; that democratic political procedures *as such* embody certain desirable fundamental normative

---

with scientific theories of inquiry. This naturalist strand is of course most closely associated with C.S. Peirce, but it has been a continuous presence besides the more subjectivist brand of pragmatism that is associated with James, Dewey, or Rorty. For the distinction, and an appreciation of Peircean-style naturalism, see Philip Kitcher, *Preludes to Pragmatism* (Oxford: Oxford University Press, 2012); Cheryl J Misak, *The American Pragmatists* (Oxford: Oxford University Press, 2013).

<sup>27</sup> Jack Knight and James Johnson, *The Priority of Democracy : Political Consequences of Pragmatism* (New York, Princeton N.J.: Russell Sage Foundation ; Princeton University Press, 2011).

principles. And because of these intrinsic features, democratic systems (or any system which embodies those values) are justified *regardless of what they happen to do*.<sup>28</sup>

In contrast, I argue in this chapter that *instrumental considerations* are decisively important when evaluating democratic legitimacy. A democratic system is valuable only if it reliably produces good outcomes (at least to some degree)—and more precisely, if this reliability is *reasonably robust*. By this I mean that reasonable people with different views of the good and different causal theories about the world and their fellow citizens, should still be able to accept any claim that democracy *does* produce good outcomes under a reasonable wide range of situations. In other words, whatever my reasonable beliefs about what the ultimate good may be, and whatever my reasonable beliefs about the causal structure of the natural and social world may be, I must have a reason to trust in the *reliability* of democracy to bring about good outcomes for its authority to be justified.<sup>29</sup>

Let us call this abstract general theory of political legitimacy *robust instrumentalism*. It is instrumental, in that political procedures are evaluated on the basis of the quality of the outcomes they produce, and it is robust since it demands that any justification must hold across a wide range of possible situations and scenarios. From the perspective of this position, procedural aspects of democratic systems should therefore be evaluated consequentially, i.e. with respect to their instrumental value in getting us to these good outcomes (whatever they may be).

---

<sup>28</sup> Except, presumably, if their actions undermine the functioning or continued realization of precisely these values.

<sup>29</sup> We can already see that this is a very strict criterion of legitimacy; and I expect that the reader may already doubt that it is possible to justify political authority on this basis. The next three chapters are given to the argument how, on the contrary, one can construct such a justification.

This *robust instrumentalism* also gives us a critical perspective on existing political systems, since this allows for the idea that *some (apparently fair) democratic procedures are legitimate while others may not be*. Concretely, this means that in situations of *conflict* of principles, the epistemic criterion should win out. Here is a concrete example: if epistemic reliability, for instance, demands a departure from strict political equality—for instance, by prohibiting Neo-Nazi parties from standing for election—this may outweigh the normative force of the principle of equality, under which such a prohibition of Neo-Nazi parties cannot be legitimately made. If the prohibition will increase the likelihood of right decisions, that is entirely sufficient to justify it.

To make the argument for *robust instrumentalism* as a fundamental theory of political authority and of democratic authority in particular, I proceed as follows. First, I map the landscape of theories of democratic legitimacy in general, and what really is at stake here. After all, nowadays opponents of democratic rule are few and far between—and even blatant authoritarians often feel the need to justify their rule with apparently democratic features (The official name of the “Democratic People’s Republic of Korea” testifies to this phenomenon). As I explain, however, it does matter *why* we think democracy valuable. Second, I argue that exclusive intrinsic political justification, what is most often called *pure proceduralism*, fails. Third, I argue that *hybrid principles* that admit a lexically secondary value to epistemic considerations fail as well: one has to either make pure procedural arguments, or admit that epistemic considerations may in principle trump procedural ones. Fourth, I consider what is often taken as a decisive argument against instrumental justifications of democracy: the empirical fact of *deep disagreement* or *value pluralism*. The view that we (as a people) typically do not agree what the right thing to do

would be, seems to favor democratic decision principles as default, or perhaps as a *residual value*. This part challenges that idea. Fifth, however, I concede that disagreement, or more generally, *deep uncertainty* about what to do does lead to the dilemma already mentioned above. However, as I argue, this does not defeat the normative attraction of an instrumental theory of democracy, it just means that the dilemma has to be resolved. The next chapter, chapter 3, elaborates on this idea of deep uncertainty and proposes a Deweyan pragmatic view of democratic legitimacy under these realistic conditions.

## 2. Justifying Rule; Democratic and Otherwise

Should we encourage more participation in our democratic system? Or should we encourage people to inform themselves about political matters and allow participation only on that condition? What if they have strongly illiberal or intolerant views? For that matter, should we perhaps introduce a qualification threshold for voters, much like we already have an age threshold for active and passive political rights? Should we give more weight to the votes of the educated, like J.S. Mill famously advocated, or should we give more weight to the votes of parents of young families, given that they, and the children in their care, will have to live with the consequences of political action (such as mounting debt, for instance) much longer than the elderly? Should we exclude extremist parties (such as Neo-Nazis) from standing for office, or is such a move anti-democratic and therefore illegitimate? Should we encourage (institutionally) the entry of new challengers—parties or candidates—into the political field, or is this unfair to established parties? Or is it the incumbent advantage that is unfair?

We could argue indefinitely which of these institutional rules would be “more democratic.” That is of course a semantic issue; and there are many ways to specify “democratic” as is convenient for such purposes. For now I will use the term “democracy” as a purely formal concept, without any normative connotations.<sup>30</sup> Democracy as it is used here is “merely” an umbrella term for certain *procedures* for making collective decisions: a set of decision rules by which a group of people can make binding decisions. There are many different variants of democratic rule, as the foregoing paragraph has already hinted at: representative, constitutional, majoritarian, unanimous, direct, deliberative, participatory, and so on. I will not here conduct a search for the conceptual “essence” of democracy—this is a task better left to others. However, for my purposes I count as democratic those decision rules that are *radically inclusive* (i.e. *everyone* within the demos gets the right to partake in the decision-making process, if not necessarily directly, then at least at some level), that allow *public contestation* of candidates, policies or ideas,<sup>31</sup> and that are broadly *responsive* to the (express) wishes of the population.<sup>32</sup> These conditions allow of degrees: a procedure can be more or less responsive, and can therefore be more or less democratic with respect to that dimension. This leaves open a wide spectrum of specific democratic procedures; and it leaves open which concrete *decision mechanisms* are created by the democratic rules of the game. This is deliberate since the purpose is to find out which of these mechanisms are indeed

---

<sup>30</sup> In chapter 5, however, I will come down on the side of a more substantive conception of democracy.

<sup>31</sup> I take this publicity to be an essential feature of democracy. Having a majority vote without any public contestation of the options on the table would mean that the choice is merely a sham.

<sup>32</sup> Here we might distinguish *responsiveness* from *congruence*. The latter refers to whether the preferences of the population are the same as those of the representatives at a given point in time; whereas the former refers to how representatives *react* to changes or shifts in the preferences of the population. A democratic system should be both; for simplicity’s sake I will just use “responsive” from here on.



justifiable. For example, majority rule is one mechanism, deliberative contestation is another. It also sharply separates the *structure* of democratic procedures from the question of the *value* of democratic decision-making. Democracy is a family of *procedural rules* and whether those procedures are good or bad is a separate question; a question that depends on our normative *standards* of good and bad.<sup>33</sup> The remainder of this chapter is devoted to clarifying these standards, without assuming that we are already committed to democracy, substantially speaking. Only in that way can we gain a critical understanding of democratic decision-making in the first place.

So what do we think is a good political system? There are of course many things we like, and people generally don't agree about what it is that they want from a political system. With varying intensity, we believe that it should be fair, treat people equally and with respect, guarantee just *and* efficient outcomes, maintain peace and increase happiness, resolve or defuse interpersonal, intergroup and international conflicts, respect human and/or liberal rights, maintain tradition and cultural achievements, foster community spirit and enable economic growth through market systems and trade. And, no question, all of these might be worthy goals. Unfortunately it is not typically true that all good things go together, and the fact that we want our political system to do or be all these things does not give any guidance as to what we should do about the institutional questions raised above.

---

<sup>33</sup> Many scholars tend to propose a moralized version of democracy, for instance that democracy is *defined* as a regime that is normatively desirable in all aspects, but this seems to me to beg the question of whether democracy is a good thing to have in the first place. I tend to think that all good things tend not to go together (at least not necessarily), and that the interesting question is what happens when fundamental values conflict.

The question facing the political philosopher is therefore: is there some more universal and more fundamental normative criterion that makes political systems desirable? Are there features of political systems that are universally valuable, despite disagreement about the ends of politics?

### *Justificatory Arguments*

A useful way to frame this question is from the perspective of the dissenter, as a question of the *justification* of (legitimate) political rule. Politics differs from other social systems in that it issues authoritative orders that are, if necessary, coercively enforced. In other words, there is not usually a voluntary choice whether to submit to political rule. As Hume famously pointed out in a criticism of Lockean consent theory, leaving the authority of a state is not usually a feasible option, much like the case of the shanghaied sailor whose option to leave the ship by jumping overboard is not a feasible one.<sup>34</sup> Contrast this with other social environments: religions, even economies to some extent admit *exit* or at least disengagement. Membership is voluntary, to varying degrees. Being subject to political rule and political authority, on the other hand, is seldom *actively* voluntary.<sup>35</sup>

Therefore political rule is in need of justification; and the justification of political rule through the authoritative system X will involve reference to a normatively attractive feature of system X. The question is, which (set of) normatively attractive features of

---

<sup>34</sup> David Hume, "Of the Original Contract," in *Essays, Moral, Political and Literary*, ed. Eugene F Miller (Indianapolis: Liberty Fund, 1987).

<sup>35</sup> The obvious exception are of course naturalized citizens.

system X are sufficient to ground the legitimacy of political authority, despite de facto disagreement about what they are? What could we point to when we want to justify coercive enforcement of political decisions to people who don't agree with that decision in the first place? As already pointed out in the Introduction, for the purposes of this dissertation I assume this "public reason liberalism," or "justificatory liberalism" as the basic theory of political legitimacy<sup>36</sup>

The first thing to note is that in this project, I am trying to look primarily at the legitimacy of different political *procedures*, rather than the legitimacy of primarily individual *decisions*; therefore we are already in the area of *proceduralist* theories of political legitimacy generally. A decision-centric theory would judge a procedure justified only when and to the extent that it makes right (in whatever sense) decisions, and non-justified (and perhaps meriting resistance) in every instance it does not. This is what Beitz and Estlund call "correctness theories" of legitimacy.<sup>37</sup> In contrast, a procedural theory allows that a procedure might be justified even though on occasion it might make wrong decisions. Accordingly, a decision that happens to be wrong can also be a legitimate one, if the procedure through which it was decided is *generally* reliable.

On which basis, therefore, could we justify a political procedure such that all its decisions are lent legitimacy? As already mentioned above, there are two basic types of such arguments: *intrinsic* and *instrumental* ones. Essentially, intrinsic theories argue that

---

<sup>36</sup> This framing of the question of political legitimacy as the question of justification goes back to Rawls. See of course John Rawls, *Political Liberalism*, John Dewey Essays in Philosophy (New York: Columbia University Press, 1993). It has probably become the dominant framework for talking about political legitimacy for contemporary diverse and pluralistic societies. See Gerald F. Gaus, *The Order of Public Reason : A Theory of Freedom and Morality in a Diverse and Bounded World* (New York: Cambridge University Press, 2011).

<sup>37</sup> Beitz, *Political Equality*; Estlund, *Democratic Authority: a Philosophical Framework*.

there is a procedural principle that demands that legitimate political decisions must be made in a certain way. Typically, such justifications take one aspect of democratic decision-making, such as its radical inclusiveness, majoritarianism, or the “one-man-one-vote” principle and argue that it reflects some underlying normative value: like fairness, equality, individual or collective autonomy, or respect for individuality. In the language of justification: a decision is legitimate if it has been reached through a procedure that reflects the appropriate normative value—regardless of what the content of the decision will be in the end, and irrespective of what the eventual decisions are like. To put it more formally:

- (1) *Principle of Justification:* A political procedure is justified if and only if it conforms to the set of formal criteria *S*.
- (2) *Factual Claim:* Democratic procedures uniquely conform to *S*.
- (3) *Conclusion:* Democratic rule is justified.

In contrast, *instrumental* arguments focus on the content of political decisions, or more generally, on the consequences of the political decision process. We might distinguish two different variants here: one that concentrates on the quality of the consequences of political *decisions*. We might call this *Instrumental Proceduralism*.

The most well-known version of this theory is *Epistemic Proceduralism*, according to which a decision is justified if and only if it has been reached by a procedure that *tends to make (objectively) good decisions*— even though we cannot be sure that every given decision we are faced with will necessarily be right. If we have reason to trust the *general epistemic reliability* of the procedure, it can be justified on that basis. David Estlund is probably the

most well-known proponent of this theory,<sup>38</sup> but as I will discuss later in this chapter, this idea has some practical problems as well.

More specifically, in order to make a convincing instrumental-procedural argument, the procedure has to reach a certain level of reliability relative to its alternatives, or ideally be optimal. Therefore we might judge a *procedure* justified (and thereby also the decisions issuing from it) if we can reasonably expect it to be best among all possible (or feasible) procedures at getting things right, even though in some instances it might not do so. More formally, an instrumental proceduralist argument might look like this:

- (1) *Principle of Justification:* Those institutions are politically justified that are elements of reference class C, and are better than all other elements of C at producing good outcomes.<sup>39</sup>
- (2) *Reliability Claim:* Democratic procedures are better than all other elements of C at producing good outcomes.
- (3) *Conclusion:* Democratic rule is justified.

---

<sup>38</sup> Although he combines epistemic proceduralism with an intrinsic criterion, making his theory essentially a hybrid.

<sup>39</sup> For example, we might think, like David Estlund, that the reference class with which to compare our proposed institutional setup is “all institutions that are subject to qualified acceptability.” David M. Estlund, *Democratic Authority : A Philosophical Framework*. In this case our reference class would contain all those institutions. If we choose a threshold standard—for example, if following Winston Churchill’s famous quip, we think that democracy is justified because it is a better form of rule than “all those other forms that have been tried from time to time,” but not necessarily the best of all possible institutional setups, the relevant reference class contains all those decision structures that have so far been tried.

### 3. What's At Stake?

Now, the reader may wonder: Why argue about *why* democracy is justified, if we already agree *that* it is justified? And what is even at stake between these two positions? Indeed, there is nothing in principle that speaks against trying to fulfill both intrinsic and instrumental standards, if possible. The reason becomes clearer when we look at cases of democratic rule where these values come into conflict. There might be a certain democratic procedure that formally satisfies the set of procedural criteria  $S$ , but which does not have the required epistemic reliability, and vice versa. Consider the following pair of examples:

Example 1: Assume an intrinsic theory requires that everyone should have equal opportunity to exert political influence. Then assume a perfectly fair voting system that nevertheless consistently produces suboptimal outcomes—Imagine a society with a permanent ethnic minority the members of whom nevertheless have equal votes and are in no *de jure* way disadvantaged within the political system, except for the fact that they are in the minority. As a consequence, however, political decisions tend to produce highly unequal outcomes, *de facto* disadvantaging the ethnic minority. For instance, anti-discrimination legislation never gets passed.

Example 2: A voting system that bars certain groups from standing for office nevertheless tends to produce the right outcomes. Imagine a system where racist parties are deemed not eligible for election and candidates may not campaign on racist terms. The prohibition is not attached to their person: they may run for office, only not on racist platforms. This system clearly violates political equality, barring as it does certain

opinions. Nevertheless, assume that this system produces less xenophobic legislation than its alternatives.

How to assess these cases? In Example 1, an intrinsic theorist presumably would find nothing wrong with the political procedure, and by analogy the decisions it reaches are legitimate. From an *instrumental* perspective, however, the reason that decisions systematically lead to unjust situations is a reason to dispute the justification of the system—and a reason to depart from strict equality in this case: perhaps anti-discrimination legislation could be mandated against majority wishes (by a Supreme Court?), or perhaps the votes of the ethnic minority ought to be given more weight (for instance by redistricting such that more candidates from urban districts are represented)? Or perhaps the most xenophobic candidates ought to be barred from standing for election, thus changing the political debate somewhat? In any case, on the basis of *instrumental proceduralism* there would be a strong normative reason to change the political system.

In contrast, Example 2 seems quite normatively deficient from an *intrinsic* perspective: indeed, its decisions are potentially illegitimate since a defined group is barred from getting involved in politics unless they change their political demands (additionally, people who would like to vote for that position are deprived of the ability to vote for this particular option). Accordingly, there would be a strong normative imperative to change the political system. From an *instrumental* perspective, however, there is nothing obviously wrong with it. Whether or not it the procedure is less egalitarian than otherwise, if banning certain positions from the political campaign spectrum leads to more just outcomes, we should do it. Political equality, in this view, is not the most

fundamental value of politics, good outcomes are (which may include *actual*, i.e. non-political, equality). Political equality is a derivative value.

These are only stylized examples—for one thing, we don't usually know with certainty which procedure would lead to which legislation, and neither is there full agreement that anti-discrimination legislation is objectively just. Nevertheless this illustrates what is at stake here: the different conceptions of political justification give markedly different judgments as to when and why political systems are deficient, and especially also *what should be done about that*.

Now that the terms are defined, the rest of the chapter argues that any plausible theory of political legitimacy *has to include* outcome considerations, and in a way that may override procedural considerations. As such, I argue against *pure proceduralist* theories, that afford normative relevance exclusively to intrinsic considerations. In other words, the chapter aims to establish that the legitimacy of a political system depends to a crucial extent on its *reliability* with which it tends to make good decisions, where that is defined in roughly consequentialist terms.

It matters what our political system manages to do, or at least what we can reasonably expect from its performance: its ability to deal with the complexities of the political world and to find solutions to difficult problems. If intrinsic proceduralism were valid, then this would not (or not primarily) matter: then first and foremost we should try to make our political system more fair or equal, regardless of whether this affects the performance. As I shall show in the following sections, that position should be rejected.



#### 4. The Indeterminacy of Pure Proceduralism

Pure proceduralism is the position that political systems ought to be evaluated only *intrinsically*, i.e. on the basis of the values that are embodied in their procedures rather than on their *instrumental* value. That is, democratic decisions are valuable (they can command legitimacy) because they are democratic, not because they are valuable for some procedure-independent reason (such as the tendency to produce good outcomes).<sup>40</sup>

This is not a self-evident claim. Why should the legitimacy of a political decision not depend on the content of the decision itself? To give a pure procedural account would require some explanation of how procedural features can lend legitimacy to a decision, regardless of what the content of the decision is. This reason, typically, refers to some more fundamental value—say, fairness, equality, or collective autonomy—that is embodied by the relevant democratic procedure. A *pure proceduralist* justification of *democracy* has to establish two propositions: (1) That democracy *uniquely* or *optimally* realizes this fundamental value, and (2) that this realization does not depend on the *outcomes*, i.e. the instrumental value of the procedure, but purely on its intrinsic features.

I argue that trying to establish these points, the democratic pure proceduralist is caught between two problems. Firstly, if the fundamental criterion representing the normative value on which democracy is supposed to be based is purely formal, it will be indeterminate which procedure actually realizes that value. Loosely defined procedural criteria are, as Estlund puts it, “incidental” values. That illustrates that it is unlikely that formal values can be *fundamental* themselves. If on the other hand the criterion is defined

---

<sup>40</sup> Of course, they may still be right for such a reason *as well*, but whether or not they are, does not matter, normatively speaking.

more “thickly,” with reference to some procedure-independent value, then its realization depends crucially on the instrumental value of democracy as well. Accordingly, pure proceduralist justifications of democracy cannot plausibly be made independently of instrumentalist considerations. Proceduralist justifications either implicitly include an instrumental element, or they are too “empty” to be of use to differentiate political systems at all.

In my discussion of pure proceduralism, I will mainly use the example of Thomas Christiano’s version of this view as the target of my arguments. There are of course many intrinsic procedural theories.<sup>41</sup> However, Christiano’s version is particularly comprehensive, as it addresses a number of arguments and fundamental considerations that are also used by other authors. It is also a particularly persuasive and therefore challenging version of pure proceduralism.<sup>42</sup> My arguments, however, apply to intrinsic-procedural theories more generally.

Let me first consider the argument that political legitimacy demands that decisions have to be made in a *fair* manner, and that democracy is valuable because it is uniquely fair. We can see which aspect of democracy is morally relevant here: democratic procedures involve a radical inclusiveness—meaning that *anyone* can influence political decisions, and no-one is barred from participating.<sup>43</sup> Furthermore, the procedure of

---

<sup>41</sup> See for instance Beitz, *Political Equality*; Jeremy Waldron, *Law and Disagreement* (Oxford: Oxford University Press, 2001); Thomas Christiano, *The Constitution of Equality : Democratic Authority and Its Limits* (Oxford ; New York: Oxford University Press, 2008); Nadia Urbinati and Maria Paula Saffon, “Procedural Democracy, the Bulwark of Equal Liberty,” *Political Theory* (2013).

<sup>42</sup> The most comprehensive formulation of this is found in Thomas Christiano, *The Constitution of Equality : Democratic Authority and Its Limits* (Oxford ; New York: Oxford University Press, 2008), especially Chapter 3.

<sup>43</sup> With the exceptions of the young and those convicted of crimes (typically).

majority decision—which is a key element of most democracies<sup>44</sup>—give *equal* formal political influence to everyone: no one person’s vote, be it in a referendum or for a candidate to become a representative, counts more than any other person’s.

*In what sense is democracy fair?*

These features, inclusiveness and formal equality of political power, ensure that no-one gets any undue advantage when political decisions are made. This kind of impartiality or anonymity seems to be at the basis of the concept of fairness.<sup>45</sup> Fairness is usually defined as a disregard of, or a “blindness to” morally irrelevant features of individuals—when choosing a certain distribution pattern, or when designing a political decision procedure. However, as a number of commentators have pointed out, inclusiveness and majoritarianism are not the only fair political procedures, and indeed they are quite probably not the fairest of them all.<sup>46</sup>

For one thing, a procedure of majoritarian voting implicitly favors the candidates who are more rhetorically gifted or skilled at presenting themselves, in short, candidates (or policies advocated by those candidates) who are more popular. It is sensitive to people’s interests, opinions and views (including views of one another). It lies in the nature of majority rule that those who are in the minority never rule, just because they hold opinions that are not shared by enough others. Minority supporters’ votes, as it were, are

---

<sup>44</sup> Although for the argument that democracy is conceptually distinct from majoritarianism, see Ben Saunders, “Democracy, Political Equality, and Majority Rule,” *Ethics* 121 (2010): 148–177.

<sup>45</sup> It is, for instance, one of Arrow’s four desiderata of a good decision procedure.

<sup>46</sup> Estlund, *Democratic Authority: a Philosophical Framework*; Anderson, “The Epistemology of Democracy;” Elizabeth Anderson, “An Epistemic Defense of Democracy: David Estlund’s Democratic Authority,” *Episteme* 5.01 (2008): 129–139; Saunders, “Democracy, Political Equality, and Majority Rule.”

almost *never* decisive. This problem is aggravated by the fact that popularity can easily be influenced by monetary resources, as any observer of American political campaigning (not only since *Citizens United v. Federal Election Commission*) can attest to.

Indeed, the possibility that there is a permanent minority that never has a chance of being decisive, is a very real problem for majoritarianism. Firstly, it happens all the time in existing democracies—one reason why we often do not realize this is that the minorities who never get their way are small and normally on the fringes of society. While there may be a regular periodic change of the guard between the big parties, Communists, libertarian anarchists, radical pacifists, religious fundamentalists, radical animal rights supporters; these people never get to be decisive in typical Western society. To me this seems *pro tanto* unfair at least in *some* sense. Now the defender of majoritarianism may say: they have an equal chance as everyone else to campaign and attempt to convince people of their cause, until eventually they may have convinced a majority; so it is not true that they never get to be decisive. However, there is nothing inherently *democratic* about that. Allowing everyone to try and convince others is compatible with any liberal form of government. Indeed, in a dictatorship they would only have to convince one person, whereas turning a *majority of all voters* over to your cause seems much more difficult. From an instrumental conception of democracy, however, this is no problem. The fact that these fringe groups never win may be *unfair*, but from the perspective of the consequences, it may well be a good thing—after all, there is a reason why these groups are called “fringe.” This already suggests that pure fairness may not be all we care about.

On this basis, for instance, Ben Saunders argues that on the basis of fairness, a procedure he calls *lottery voting* would be preferable. Here, everyone votes for her favored policy or candidate, and the final decision is made by randomly drawing one of the votes. This is more fair than majority decision since it gives those holding unpopular opinions a non-zero chance of winning—albeit a chance proportionally smaller than that of supporters of more popular proposals.<sup>47</sup> Arguably, this corresponds more closely to the ideal of fairness: since every citizen’s vote has an equal chance of being decisive in a given case, no-one has an ex-ante advantage based on popularity.

Going even further, Estlund has argued that if we are only concerned with procedural fairness, we should not even have a random draw of votes, but a random selection from the universe of possible alternatives.<sup>48</sup> Only a truly random choice (such as a coin flip or a roll of a 20-sided die if there should be more options) will ensure that no feature can give any of the options an undue advantage. Lottery voting is fair on the level of individuals (since every individual has an equal chance of being decisive), but it is not fair between options—the more popular ones have a higher chance of winning. Therefore, a completely random choice across options is presumably the most fair procedure. Estlund considers this a *reductio ad absurdum*, showing that fairness is at best an incidental value, but cannot be a fundamental one in itself. The point is not to show that a purely random choice is necessarily normatively more attractive than a majority decision. Rather, this shows that when we want to reject the purely random choice, it must be on the basis of some other external value; it cannot be because it is unfair. If that

---

<sup>47</sup> *ibid.*; Ben Saunders, “Combining Lotteries and Voting,” *Politics, Philosophy and Economics* 11.4 (2012): 347–351.

<sup>48</sup> Estlund, *Democratic Authority: a Philosophical Framework* 82.

is the only thing we care about, why not a random choice—and the converse: *if we think that a random choice is not justifiable even though it is fair, why is that so?*

Now we might think, with Christiano, that what this shows is that equating fairness with complete disinterestedness is misunderstanding the concept. When we consider the fairness of a procedure, we do not think it should disregard *all* features of individuals.<sup>49</sup> Rather, fairness is context-dependent; what we think fairness involves depends on the task at hand. Fairness demands insensitivity not to *all* personal features of individuals, but only to features *irrelevant* to the purpose at hand. This of course introduces a procedure-independent normative criterion: that of *relevance*. There is nothing in the concept of fairness that will point us to which features are or are not relevant in a given context. This depends on the context and the purpose of the procedure. This attention to what the procedure is supposed to achieve might already look like a major concession to instrumentalism. However, Christiano argues that which type of fairness is required in the political sphere is relatively obvious.<sup>50</sup> A fair beauty-contest should disregard all features of candidates but their looks, and the losers should not complain of being treated unfairly unless they were judged on features irrelevant to the contest at hand (e.g. their personality). Committing a foul on the basketball court is unfair, being tall and shooting over the head of your opponent is not. Similarly, Christiano argues, we might think that sensitivity to people's wishes and opinions while disregarding all other incidental features is the kind of fairness we want in the political sphere.

---

<sup>49</sup> Thomas Christiano, "Debate: Estlund on Democratic Authority," *The Journal of Political Philosophy* 17.2 (2009): 228–240.

<sup>50</sup> *ibid.*, 232.

However, it is far from obvious that this is the only kind of fairness we want from our political system. Why, in the political context, should a popularity contest—which is essentially what majoritarianism is—be the unique fair way to organize things? How about a fair meritocratic procedure? Consider a competitive examination testing potential rulers’ knowledge and intelligence that is open to everyone and whose evaluation criteria are publicly known. This system—under ideal conditions—is sensitive to candidates’ competence, and disregards features such as popularity and rhetorical skill. Provided the exam system was not subject to manipulation, and provided that everyone had an equal chance to gain the necessary qualification, I do not see how someone scoring low and therefore missing out on political power could complain of being treated *unfairly* in such a system. Indeed, in its objectivity it seems, if anything, *more* fair than the popularity contest system. Imagine, for instance, a school where the valedictorian is determined not by performance on objective tests, but by a vote among the students. Both are, arguably, fair procedures in some sense; and there is nothing, at least so it seems to me, about the political realm that “naturally” favors that particular manifestation of fairness. What this means is that if we still think democracy is to be preferred, we have to go beyond a formal value like fairness to give a reason why it may be justified.

## 5. Substantive Values Coming In

In other words, the procedural advantage of democracy has to be based on some more substantial value than fairness; a “thicker” reason to prefer the popularity contest to the competitive exam or the random choice across possible options. The concept of

“procedural fairness” in itself leaves open many different possible ways to interpret it. The same goes for other purely formal values, such as “equality.” This means that in order to give a plausible justification of democracy, we have to commit to a more substantive standard. We may try and give a substantive standard that is relatively uncontroversial, but it remains a substantive standard. To quote David Estlund on this point:

The selling point is no longer said to be that fair proceduralism steers clear of all procedure-independent standards. The point of insisting on the importance of procedural fairness is now to emphasize that this attention to voter preferences is a very minimal and uncontroversial substantive standard for outcomes. Approaches to democracy that rely on more robust standards for just outcomes such as principles of justice, or experts or procedures to guide us to good outcomes, are criticized by fair proceduralism not for importing procedure-independent standards, but for importing standards that are too controversial.<sup>51</sup>

This means that we have to pay attention to the *reason* given why, for instance, we should prefer the democratic to the meritocratic version of fairness. What could this substantive value be, that is nevertheless consistent with the two requirements of pure proceduralism? It should justify the use of democratic procedures, while *denying* that the outcome of the procedure is normatively relevant. The value most often cited here is the *equal moral status* of each individual, which means everyone deserves to be treated with *equal respect* for their moral agency. Christiano for instance argues that political legitimacy requires the *public affirmation of equal respect* for each individual, and that deciding things democratically uniquely satisfies this demand.<sup>52</sup>

I wonder, however, whether the realization of such a value does not at all depend on the *content* of the actual decisions, but only on the procedure by which they are

---

<sup>51</sup> Estlund, *Democratic Authority: a Philosophical Framework* 83.

<sup>52</sup> See Christiano, *The Constitution of Equality : Democratic Authority and Its Limits*. In addition to political equality, the ideal of public equality also requires an “egalitarian baseline” of liberal rights and an unconditional basic income.



reached. Does equal treatment in this sense not require something *more* than merely an equal vote? That is, does it not matter what the *result* of using the procedure is? Indeed, we might think that equality of respect for different people's interests and agency may even require a *departure* from majoritarianism or radical inclusiveness. This depends on what we believe equal respect really entails in terms of political principles. I shall just briefly discuss a few ways such an argument could be made.

As Ronald Dworkin has pointed out, we can understand the idea of "political equality" in a multitude of different ways, depending on what exactly it is that should be equalized.<sup>53</sup> We can equalize strictly the direct *impact* an individual has on the decision. This would obtain if everyone had one vote, they were weighted equally, and there were no differences in districting that would skew the distribution of political power. Then, on the other hand we might think that what is to be equalized is not the direct impact, but the *influence* on the political outcome. Clearly people with more resources, access to media, and perhaps those within larger coalitions have a larger influence in a system of equal votes.<sup>54</sup> Therefore, equalizing political influence might require giving those with fewer external resources *more* political impact: the votes of the poor could be weighted more heavily, or they could get veto powers over certain areas of politics.

Thus, do we think that equal respect demands equality of impact or influence? The dreaded question of *relevance* comes back in: which is the relevant kind of inequality of respect? It seems difficult to decide this without considering what either procedure would *tend to produce*. One immediately obvious argument would be that equality of influence is

---

<sup>53</sup> Ronald Dworkin, "What is Equality? Part 4: Political Equality," *University of San Francisco Law Review* 22.1 (1987): 1–30.

<sup>54</sup> Beitz, *Political Equality*.

preferable since everything else would lead to distorted outcomes: outcomes that are disproportionately influenced by the wealthy and/or charismatic. But this is at least partly an instrumental consideration; pure proceduralists want to exclude outcome considerations from influencing the judgment of the procedure.

Perhaps *political equality* generally (whether of impact or influence) does not even necessarily follow from the fundamental ideal of public realization of equal respect? We might think, for instance, that equal respect for persons requires that they have political power *proportional to their stake in the matter at hand*. This concept (we might also think of this as political *equity* instead of *equality*) has been elaborated by Harry Brighouse and Marc Fleurbaey.<sup>55</sup> It still allows for political equality as a special case: when stakes are equal, or perhaps when there is no reliable way to ascertain the difference in stakes.

But in cases where stakes *are* unequal, and the inequality can be roughly estimated, would equal respect not require unequal political influence? Why should heterosexual persons get to decide, with an equal degree of influence, whether homosexual couples should be able to get married? The stakes in this issue (despite the protestations about the “devaluation of heterosexual marriage” to the contrary) are clearly radically different. Alternatively, why should the elderly get the same political influence over education policy as young parents? The elderly have much less of a stake in the fate of current young children than their parents, let alone the children themselves. We might not go as far as to “disfranchise the elderly,” as Philippe Van Parijs’s provocative title has it, but why not give young parents some extra votes on issues of early

---

<sup>55</sup> Harry Brighouse and Marc Fleurbaey, “Democracy and Proportionality,” *The Journal of Political Philosophy* 18.2 (2010): 137–155.

childhood education.<sup>56</sup> Indeed, as a whole class, we might think that interests based on *purely other-regarding* preferences should in principle be considered to be of “lower stakes” than those based on preferences about one’s own security and well-being.

It is not obvious that political inequality in those cases would be tantamount to disrespect. Quite the opposite: equality may seem disrespectful. This clarifies another point: departures from political equality are undesirable only when they could not be justified to citizens: if the inequality were based on *invidious* ascriptions of inequality. The principle of proportionality shows that not all departures from strict equality are necessarily *invidious* across the board.

Furthermore, not only stakes, but also competence may vary. As Arneson puts it: “appropriate respect for an agent’s rational agency capacity is shown by recognizing it as what it is. It shows no wrongful disrespect to me to notice that I am imperfectly rational and to take efficient steps to prevent my proclivity to mistakes from wrongfully harming others or for that matter myself ... Respect for rational agency should not be interpreted as requiring us to pretend that anyone has more capacity than she has or to pretend that variation in capacity does not matter when it does.”<sup>57</sup>

In short, once we move beyond the purely procedural concept of fairness and ground political legitimacy in a more substantial value, like equality of *respect*, it becomes problematic to argue that a specific set of decision principles that embody political equality uniquely follow from this value. There is nothing obvious in the concept of

---

<sup>56</sup> Philippe Van Parijs, “The Disfranchisement of the Elderly, and Other Attempts to Secure Intergenerational Justice,” *Philosophy and Public Affairs* 27.4 (1998): 292–333.

<sup>57</sup> Arneson, “Democracy Is Not Intrinsically Just” 52.

fundamental equality of respect that would pick one particular conception of how to realize it over another.

Moving beyond the procedural dimension there is the following additional question: does treating someone publicly with equal respect not require anything in terms of substantial outcomes? For instance, we might think that equal respect requires respect for a person's *interests*, not only for their agency. Christiano argues that "Each person justly claims an equal share in the resources or opportunities for influencing this process of decision-making because each person has fundamental interests in being able to participate in this process."<sup>58</sup> However, it is not obvious that the fundamental interest people have in this kind of participation is the only relevant interest people might have—or even the most important one.

For instance, for someone who places a lot of value on his religious beliefs, respect might involve being able to exercise these beliefs freely; this is a *substantive* claim. Such a person might (reasonably) not be entirely satisfied when she is told that she is given a vote equal to everyone else's, but that the majority might well restrict her religious freedoms. A poor person, equally, might have a much greater interest in additional resources rather than the vote: respect might require a much more egalitarian income distribution. To be fair to Christiano, he does address these problems to some extent, in so far as he argues that the right system requires a guaranteed minimum income and a set of liberal rights; beyond that, however, for him the political system ought to be selected purely on procedural grounds.

---

<sup>58</sup> Christiano, "Debate: Estlund on Democratic Authority" 233.

In sum, political equality may be *one* way to ensure a public affirmation of the equal respect principle, but it is not clear why this fundamental principle should require precisely (and only) that, and not something more substantive. But if we admit the latter, we have already left pure proceduralism: in that case it would matter decisively what the political system *does* decide, and not only *how* it decides.

## 6. Practical Concerns

These are already some considerations that suggest that pure proceduralism alone is insufficient to establish legitimate and legitimately enforceable political authority. But there is another key reason why outcome considerations are essential. In fact I have very strong *practical* interests that political decisions should tend to be of some quality. These may even trump my interest in having a say in the political decision process. That is, we can make a good argument that *even purely from a perspective of the equal respect for everyone's interests*—positing no transcendental truth about politics—the instrumental reliability of a political procedure is crucially important.

There are two reasons: the first is that I have limited cognitive capacity, and the second is that I realistically and in practice cannot influence *all* decisions that affect me. These two points have in a similar way been elaborated by Philip Kitcher: they show that I have a fundamental interest that the political decision process be informed by an adequate system of *public knowledge*.<sup>59</sup>

---

<sup>59</sup> see Philip Kitcher, “Public Knowledge and the Difficulties of Democracy,” *Social Research* 73.4 (2006): 1–20; Philip Kitcher, *Science in a Democratic Society* (Amherst, New York: Prometheus Books, 2011).

The reason underlying this is as follows: my expressive interest in *participating* in the political decision process on equal terms in order to have equal respect accorded to me is not my most *fundamental* interest. Presumably, the most fundamental interest I have is to have an *effect* on the shape of the common world in a way that advances my interests. Imagine a process where everyone gets to participate equally (say, everyone gets the same amount of speech time before the decision), but then the votes are thrown away and the decision is taken by some entirely random procedure. This is a perfectly good public expression of equality; what is wrong with this picture is that the participation of everyone is not actually causally connected to the outcome. I participate on equal terms, and enjoy procedural respect in that sense, but I am not influencing the outcome at all. Despite the equal participation the political process does not *realize* anyone's interests (except possibly by accident).

Therefore, I have a fundamental interest in the political process *realizing* my interests, subject to everyone else having the same interest. In Kitcher's words, we want to have an appropriate level of *control* over the policy process, not only symbolic equality. But if we admit that it is not the expressive nature of equal political participation but the appropriate causal control I have over the outcome that matters, we have to look at how this connection may break down.

In practical terms, there are many judgments that are beyond my cognitive capacity. In what way a given policy proposal would affect me and my interests is in many cases obviously beyond my understanding—and indeed beyond anyone's individual understanding. Often I cannot figure out how my interests would be best advanced in the complex situation of contemporary mass society, where everyone is trying to realize their

goals: “If the voters cannot see what is happening, or if they are unable to trace the confinements they feel to their source, their votes are unlikely to serve as a means of control.”<sup>60</sup>

If that is so, however, I have a *practical* interest that decisions taken through the political system be *good* in some sense; an interest that goes beyond an interest in my personal participation. In cases where figuring out how to realize my interests in the complex system that is the common world is beyond my (and anyone’s) capacities, the *reliability* of the decision-making procedure becomes important.

Consider this example: assume a large majority of citizens is in favour of minimizing the number of violent deaths in society.<sup>61</sup> Assume further that the death penalty actually has no measurable deterrent effect on the number of murders, and that this is known to criminologists. Finally, assume that a majority believes that the death penalty has this deterrent effect, and that executing convicted murderers will minimize the occurrence of violent death.<sup>62</sup>

In this case, presumably, people’s interests would be best served if there was a majority vote on the normative goal of politics, but the means to get there would be determined in some other way (or by someone else). In this simplistic cause-and-effect model, voting separately on ends and means would mean the murder rate would not fall, and a lot of people would be needlessly put to death—and the same goes if there should be a direct vote on the death penalty.

---

<sup>60</sup> *ibid.*, 78.

<sup>61</sup> I would think it’s realistic that there is near-universal agreement on this

<sup>62</sup> Despite the fact that the executions are themselves violent deaths, the majority believe that overall, this reduces the number of deaths.

Now the pure proceduralist might reply and argue that people have a fundamental interest in directly selecting the policies that govern them; therefore the fact that they are wrong on the factual question does not matter—what matters is that they themselves get to choose. But it is not obvious why this agency interest citizens have is more important than the interest they have in seeing their views realized in the world. Nobody wanted that outcome; note that it is not the case that people’s opinions somehow do not reflect their “deep” interests or that they suffer from false consciousness; quite the contrary: their opinions do reflect their interests, it is only that their factual beliefs are wrong.<sup>63</sup> While the “right to make mistakes” might be a part of the value of autonomy, a system where such failure is systemic is not salvaged by its procedural virtues. It would not be justifiable because it is epistemically unreliable; and we can make such a judgment purely on the basis of equal respect for individuals’ interests, without assuming a transcendental standard of political truth.

Furthermore, for practical reasons, it is infeasible for me to personally exercise effective control over every aspect of the political sphere that affects me. As Kitcher puts it again: “Within contemporary industrial and postindustrial democracies, however, the idea of public control of decisions that affect all citizens looks ludicrous. An intricate division of labor means the life of any individual is affected by the actions of vast numbers of others, so enormous numbers of institutional mechanisms need to be in place to constrain interactions taking place at unfathomable distances [...] You are in no position

---

<sup>63</sup> A different problem that arises from the possibility of decomposing political decisions into several parts (ends and means, or different parts in some other sense) has become known as the *Doctrinal Paradox* or the *Discursive Dilemma*. Here, the outcome of a complex decision is determined by the way in which the decision is partitioned, and whether they are decided separately or “as a package.” It is explored in several works. For instance, see Christian List and Philip Pettit, *Group Agency : the Possibility, Design, and Status of Corporate Agents* (Oxford: Oxford University Press, 2011).



to assess all these consequences or even to recognize some of them.”<sup>64</sup> We cannot spend all our time thinking about political questions, and even if we did, it is obviously impossible for everyone to *directly* influence the political process all the time. Therefore, realistically, there must be substantial parts of the political territory that are unknown to me. The idea that I could therefore exercise democratic control over all aspects that affect me is not viable; and I have a fundamental interest that those issues which cannot personally control be resolved in the best way.

Note that this point does not apply only to systems of direct democracy—if it did, it would not be so interesting. Representative democracy with professional, full-time, elected politicians equally suffers from this problem: the complexity of the political sphere does not go away just because the options are now reduced to a smaller number of candidates or parties. It just means that the choice of which package of policies to vote for becomes more complicated: determining which candidate will really realize my interest is quite a complex issue, as I have to estimate what the candidate will do with respect to all issues that affect me, *and* whether that decision will be good. That is as unrealistic as trying to actively decide on all issues myself. If I use heuristics to choose candidates (e.g. “The Democrats will generally act in my interest”), then it becomes obvious that what Democrats *actually do* is important for me.

Therefore, to the extent that I *can* exercise real control over the policy process through my participation, I have a fundamental interest that this control be effective; and to the extent that I cannot realistically exercise control, I have an interest that the process that determines those issues have adequate *reliability*.

---

<sup>64</sup> Kitcher, *Science in a Democratic Society* 78-79.

This shows that even *on their own terms* pure proceduralist theories, when analyzed realistically, are insufficient. Unless we want to postulate a radically simpler world than the world we find outside, epistemic considerations *must* come into play. In order to justify political authority to me, it is plainly not enough to just say that I can participate on equal terms in the political process.

## 7. Intermediate Summary

What this chapter has aimed to establish so far is firstly, that pure proceduralist arguments that are based on purely formal values (like fairness) do not uniquely pick out democratic procedures: lots of procedures are fair in some sense, and clearly, not all of them are equally acceptable. Which sense is the *right* one to justify political authority has to be determined from a more substantive standpoint; the question of *which* fair procedure is justifiable cannot be answered from the perspective of that formal value itself. Instead, there must be a reference to some substantive standard to tell us why we want a fair test of competence for medical practitioners (we want to find the most competent doctors), a fair beauty pageant (we are looking for the most beautiful person), or a fair game of basketball (we want to find the best athlete).

Secondly, however, once we try and base our pure proceduralist argument on a more substantive value—as an example I discussed Christiano’s idea of public realization of equal respect for individuals—it seems that *outcome* considerations become ever more important. I have shown ways in which they not only matter, but indeed may outweigh procedural considerations altogether. Firstly, it is not obvious that equal respect does not

require anything in terms of *substantive* policy outcomes; and secondly, I have good *practical* reasons to demand epistemic reliability from the political process: because of the complexity of political questions. Thus, equal respect for individuals' interests depends crucially also on the outcomes of the process.

What this amounts to is to show that pure procedural arguments are by themselves insufficient to ground political authority: when we evaluate a political system, we do and should care about what it does. Even if we care only about equality or fairness (or perhaps especially when we do so) the substance of the decisions made by a political system matter. More precisely, political legitimacy depends on us somehow being able to show that our political system does produce good outcomes, at least to some extent.

## 8. Against Hybrid Principles.

We are still trying to find convincing arguments to justify the exercise of political authority, and democratic authority in particular. Arguments based purely on proceduralist considerations were found to be lacking, since they either fail to uniquely pick out democratic principles (in that case it would be indeterminate), or have to introduce some outcome considerations to remain plausible.

But the intrinsic theorist might relax the call for “purity” of the procedural principle, and admit that epistemic considerations play at least a secondary role. In the previous section I have argued that equality of respect may require attention to a procedure's outcomes, so in response the pure proceduralist may argue that equality of

respect *first* entails some procedural requirement, and then *second*, epistemic considerations come into play when we choose among the procedures allowed by the first criterion. This would amount to a *hybrid argument* of the following sort:

- (1) Only procedures that satisfy procedural criterion C are potentially justifiable.
- (2) Among those procedures, only the one(s) that produce(s) the best outcomes is/are justified.
- (3) Democratic procedures satisfy criterion C and are most epistemically reliable among those that do so.
- (4) Democratic procedures are uniquely justified.

In the end, this is the kind of argument Estlund ends up making in *Democratic Authority*, which is why at this point I part ways with him. In his argument, the criterion C is the so-called “qualified acceptability criterion,” which states that only procedures whose epistemic superiority is potentially acceptable to all qualified points of view are justifiable.<sup>65</sup> What this means in the context of Estlund’s argument is that a claim of reliability has to be acceptable to all qualified points of view; and in his theory this acceptability ends up being defined procedurally. In particular, a claim to epistemic reliability is acceptable if it does not employ *invidious comparisons* of the differential epistemic capacities of individuals. Since for him all of these interpersonal comparisons of epistemic competence can potentially be seen as invidious, this means that *only those* procedures are justifiable that give the same level of political influence (or the same chance of influencing the political outcome) to everyone; this includes majority rule

---

<sup>65</sup> Estlund, *Democratic Authority: a Philosophical Framework* ch. 3.

procedures, but also lottery voting or a simple random choice among possible options (both of which give equal—zero—influence to each individual). Finally, among the choice between random, lottery, or majoritarian egalitarian procedures, Estlund’s hope is that the latter are (or can be shown to be) the most epistemically reliable. The concern with the qualified acceptability of epistemic superiority claims therefore creates a two-tier argument that incorporates procedural and epistemic elements.

As already mentioned, “qualified acceptability” is defined *procedurally*. What is acceptable does not depend on substantive content. Epistemic considerations only come into play within the frame set by the acceptability criteria. This notion is challenged by robust instrumentalism. We need to show that acceptability can robustly be defined in outcome terms as well. This argument is addressed in the next chapter; here I only argue against these hybrid principles.

In any case, we could characterize Estlund’s conception as giving *lexical priority* to a specific procedural criterion: procedural considerations can defeat epistemic considerations, but not vice versa. Even if we had the perfect (non-egalitarian) procedure that got the right answer every time, and everybody knew it, the procedure would be illegitimate if someone could justifiably object to its use.<sup>66</sup>

Initially, this sounds plausible: Let’s say we know that some procedures would be extremely epistemically reliable (say, giving experts exclusive authority over their field of expertise), but they are not acceptable to all qualified individuals—some people might reasonably question the competence of these experts. Then it may seem reasonable,

---

<sup>66</sup> Even if they should not do so *in fact*.

according to Estlund, to restrict the potential candidates for legitimacy to democratic procedures, since they are the procedures whose epistemic optimality nobody can reasonably deny.

In order for such an argument to go through, one would have to establish the following two propositions: (1) for any non-egalitarian decision procedure, it is *in principle reasonable* to deny the epistemic reliability of that procedure; (2) for some democratic decision procedures, it is *in principle not reasonable* to deny their comparative epistemic reliability.

Unless (1) is established, Estlund's hybrid epistemic proceduralism would allow non-egalitarian decision procedures, which he wants to avoid; and unless (2) is established, the theory would remain indeterminate between all egalitarian procedures—including the random choice. The difficulty is now how to come up with a conception of “reasonable” that fits the bill. Consider the following example, drawn from Estlund's discussion of *primary bads*.<sup>67</sup> Primary bads are catastrophic events, like famine, war, economic or political collapse, or genocide; events that nearly everyone agrees are bad things, and which therefore are pretty much uncontroversially undesirable.

Here is the example: Assume that political collapse is a primary bad—this means that it must be unreasonable to deny that it is a bad. Consider a state where nearly everyone agrees that political collapse should be prevented, but that has a small radical libertarian minority that denies the legitimacy of the state outright. They are fighting (non-violently) for political collapse since they actually believe that to be beneficial. They

---

<sup>67</sup> *ibid.* ch. 9.

are not morally perverse, they just think that any state beyond the night-watchman kind is harmful and oppressive, oppressive to such a degree that collapse is to be preferred. If this state had a democratic system it would be very unlikely that there will be political collapse, as the radical libertarians are but a small group and have not had much resonance with the wider public so far. But there is a small chance, since if the radical libertarians get to participate in politics (especially if they are wealthy) they may in the end convince enough people of their goal, and once they have a blocking minority, political collapse may well ensue. Therefore, a system where they were excluded from politics would be *even more reliable*, since even the small chance of political collapse would be removed. Now I can see two problems here:

*Either* we consider their opposition to be in principle reasonable. Then we cannot exclude the libertarians from participating in politics, since they could *reasonably deny* the epistemic reliability of the ensuing system. They think that political collapse would be a good thing, so of course from their perspective they do not think that a system that will guarantee stability of the existing unjust order is “reliable” in any sense. We cannot justify the exclusionary system to them by arguing that preventing political collapse is a good, since they reasonably deny that. Even though a system that excludes them will prevent the primary bad of political collapse, it is not justifiable *to them*.

So the choice is between democracy and some random procedure. The radical libertarians prefer a random decision procedure to democracy, since that will give them a bigger chance of winning and causing political collapse than democracy. In democracy they will also have a chance, but it is a smaller one. The issue then is this: how *can* we justify the choice of democracy over a random procedure *because it is more likely to prevent the*

*primary bad of political collapse?* How can it be is *unreasonable* to deny this latter advantage of democracy over lottery voting while it being *reasonable* to deny the advantage of the exclusionary system over democracy? If we think that preventing political collapse is universally good, why is that not enough to justify democracy that excludes the group of libertarians? And if we really take the libertarians' disagreement seriously (if we hold it to be reasonable), why is it ok to impose democracy on them, on the basis of an epistemic justification? Presumably If I am a member of the majority, I like democracy and the exclusionary system, and if I am a libertarian I reject either.

So, my question would be: what is the *standard of epistemic reliability* that squares this circle? How can we reasonably assert that that democracy is epistemically reliable? We have to have a way to *independently* ascertain this in a way that everyone can accept: but if we have such a reliable and acceptable procedure, why is that procedure itself not a justified political decision mechanism?<sup>68</sup>

The problem of course derives from asserting, at the same time, an *independent* standard of epistemic reliability that grounds democracy's relative reliability among the procedures that satisfy the procedural criterion, while denying that any such independent standard may be used to justify non-egalitarian procedures. In my view, any definition of reasonableness that allows this would be somewhat *ad hoc*.

Now we can see that the intuitive attractiveness of the hybrid theory trades on a slip in perspective between the viewpoint of the reasonable citizen, and the political philosopher. Strictly speaking, only the *citizen's* acceptance is necessary and sufficient to

---

<sup>68</sup> Similar arguments to this and the following are made by David Copp, "Reasonable Acceptability and Democratic Legitimacy: Estlund's Qualified Acceptability Requirement," *Ethics* 121.2 (2011): 239–269.



justify political authority procedurally. That *the philosopher* knows what is the right thing to do is not relevant to the justification at all, unless it can be shown that it would be *in general* unreasonable to deny that independent standard.

The question is therefore what normative significance we attribute to the citizens' actual or potential disagreement about what to do. If we think it disqualifies otherwise valid epistemic justifications of non-democratic procedures—because all truth-claims can be reasonably disputed—it should disqualify instrumental defences of democracy as well. This route would allow only procedural justifications of democracy, if any. Then we would be back to square one, if all epistemic standards are subject to reasonable objection. This is one way a hybrid argument might collapse into a version of pure intrinsic proceduralism.

If on the other hand, we think reasonable disagreement does *not* entail the invalidity of all justifications based on truth-claims, then there is no reason why instrumental justifications should be subordinate to procedural considerations. If it is valid to refer to independent standards of epistemic reliability when justifying democracy, why is this restricted to procedures that pass some procedural test?

The only way to make such an argument would be to use “reasonable” in a specific—and in my view, ad hoc—sense: that it would be unreasonable, in principle, to dispute democracy's epistemic reliability, while it could be reasonable to dispute some other procedure's reliability. I cannot immediately see any basis on which some such a claim could possibly be made.

What this discussion entails is that, in a sense, a hybrid principle that assigns epistemic considerations a secondary role is “unstable.” We either have to make a purely procedural argument, or we have to go all the way and agree that epistemic considerations can trump procedural considerations when it comes to justifying rule. The important upshot is that democracy’s epistemic reliability must be demonstrated “on its own”—not only in comparison to random mechanisms, but relative to all feasible political systems. The task, as it were, is more difficult than hybrid theories make it seem.

I have already written about the indeterminacy of pure proceduralism: there seems to be no obvious way to establish a positive version of the claim that outcome considerations do not matter at all. However, more needs to be said before we can establish epistemic proceduralism: for one thing, pure proceduralism may be based on a purely negative account. We cannot agree on what to do, but we need to do something—and a democratic choice seems like the appropriate way to do this when there is such disagreement. Perhaps democracy is just a *residual value*, the default decision mechanism under conditions of disagreement.

Before we can establish epistemic proceduralism, therefore, we need to figure out what actual *disagreement* in politics entails: does it invalidate all truth-claims in the context of democratic justification, or does it entail something else? This is the subject of the next section.

## 9. Political Equality as a Residual Value

So far I have been trying to show that the instrumental or “epistemic” reliability of a political system is a necessary element in its justification. However, there is of course a problem that already came up in the previous section: how can we assert epistemic reliability of any procedure if there is disagreement about what is *right*? In modern, plural societies there is typically quite astonishing diversity of what people want, about what they believe in moral and factual matters, and what they think the purpose of politics is. And, as many observers note, this disagreement is both *extensive* and *intensive*. That is, it covers many, perhaps even all, fields of politics; but beyond that, it is *deep* disagreement, disagreement that cannot easily be rationally resolved—or perhaps not at all. Any given disagreement may go “all the way down,” i.e. there might be no fundamental shared values or beliefs to which a rational resolution could appeal. Recall the context of justification: the epistemic superiority of democracy needs to be established relative to feasible alternative systems, and this demonstration must be acceptable from all reasonable (or “qualified,” in Estlund’s terms) viewpoints.

The question therefore is: what does this fact of reasonable disagreement entail? One point of view is probably best summed up by this quotation from Jeremy Waldron: “rights-instrumentalism seems to face the difficulty that it presupposes our possession of the truth about rights in designing an authoritative procedure whose point it is to settle that very issue ... There seems, then, something question-begging about using rights-

instrumentalism as a basis for the design of political procedures among people who disagree on issues such as this.”<sup>69</sup>

This suggests that when we are looking for a procedure to settle real and deep disagreement, we cannot ground the justification of the procedure on some substantial understanding of the truth: precisely because it is disputed what that truth is. If we knew the truth already, we would not need to look for a procedure to settle disagreements.

This point is crucial: If I intend to justify political authority procedurally, to someone who disagrees with the substance of particular political decisions, it will not do to tell her that the decision is right. But we have to come to *some* conclusion: and in the absence of good arguments to the contrary, democratic (majoritarian) procedures seem like a good “default.” If my group of five friends is trying to decide which movie to watch at the theater, and no-one has a special claim (e.g. I have already seen all the movies on show but one), or an infeasible preference (e.g. horror movies make me feel sick), taking a vote seems like the “natural” way to resolve this. Following this thought, we might think that the fact of disagreement alone is a good *prima facie* reason to support democracy as a “default.”

However, I have already mentioned that we *do* seem to care about the quality of political outcomes. In order to counter this reasoning, therefore, we would need what Estlund calls a “formal” argument for the epistemic reliability of procedures—one that does not make reference to any substantial idea of the truth.<sup>70</sup> As Waldron puts it: “we

---

<sup>69</sup> Waldron, *Law and Disagreement* 253. By “rights-instrumentalism” Waldron refers to the idea that rights, including political rights that define democratic rule, have only instrumental value.

<sup>70</sup> Estlund, *Democratic Authority: a Philosophical Framework* ch. 9.

might say instead that we should choose or design political procedures that are most likely to get at the truth about rights, *whatever* that truth turns out to be.”<sup>71</sup> Beyond that, not only do we disagree about what is substantially true with respect to specific questions, we often disagree about basic epistemological questions: we don’t even agree on what makes a claim true, what would count as relevant evidence one way or the other.<sup>72</sup>

The question is: is it possible to give a *robust instrumental* justification for democratic authority, given the disagreement and plurality of values and beliefs? More specifically, could such a procedure be justified in a way that is beyond reasonable objection? If there is disagreement about the truth, as well as the *path to the truth*, how can such an argument be made? In this section I only give some initial considerations on the possibility of such a justification, while the next chapter will give a more precise argument how we can understand this.

The first issue we have to address is the *nature and moral status* of the disagreement. Firstly, presumably not all disagreement has moral force. We might think that a procedure does not need to be justified to the morally perverse or people holding beliefs mistakenly or in an unreflective and unconscious manner. If my disagreement with you would just disappear if I stopped and thought about my beliefs; or if it stemmed from my desire to harm other people (especially you), then my disagreement may not be of moral significance. This could be subsumed under the heading “unreasonable disagreement.”<sup>73</sup>

---

<sup>71</sup> Waldron, *Law and Disagreement* 253.

<sup>72</sup> Laura Valentini, “Justice, Disagreement and Democracy,” *British Journal of Political Science* 43.01 (2012): 177–199.

<sup>73</sup> Waldron, in so far as he accepts that governments can, in principle, be legitimate, is committed to the view that not *all* disagreement, let alone *all potential* disagreement can be morally relevant. See David M Estlund, “Jeremy Waldron on Law and Disagreement,” *Philosophical Studies* 99.1 (2000): 111–128.

But even so, plenty of reasonable disagreement remains even between people who have reflected and who are not out to hurt each other.

However, reasonable disagreement of course does not rule out the possibility that there is a truth about the matter. Such an inference often seems commonplace in normative theory, but in many other areas of inquiry this is of course not taken for granted. As it were, the epistemological problem of the access to the truth tends to be seen as independent of the ontological question about the existence of that truth. This is true of causal/factual beliefs but also of moral beliefs. With respect to factual beliefs about the world it is obvious that disagreement about truth does not entail non-existence of the truth: we do not take scientific disagreement to indicate that there is no objective fact that may in principle resolve the disagreement.<sup>74</sup> Speaking with policy in mind: when we disagree whether a policy A will have the intended effect B or not, this does not entail there is no right answer to this question—even though it might entail that we don't (yet) know what it may be.

With respect to normative issues as well, the disagreement entails an epistemological, not an ontological claim. Because people reasonably disagree about something does not mean that there is no right answer, it just means that there is *uncertainty* about what that answer is. Waldron, Valentini and other proponents of the residual argument for democracy seems to assume that this uncertainty is so deep as to make it impossible to ground any epistemic (instrumental) accounts of political authority,

---

<sup>74</sup> Even the most skeptical theories of science are skeptical about our ability to gain reliable knowledge about the factual truth (for instance because phenomena are indeterminate and allow for multiple explanations), not about the existence of that truth.

or more precisely, that the problem of uncertainty makes any instrumental justification immediately suspect.

It might be that one of the parties to the dispute is right, but we have no way of knowing which, so we cannot use this sense of “right” in justification. However, the assumption that there is a truth that could, in principle be discovered, remains plausible. With respect to normative questions (such as, “are the consequences B of policy A just?”) this seems more problematic than with respect to causal questions. However, the distinction between factual and normative question—or between facts and values—is less of a bright line than it would initially seem. Indeed—and I will return to this issue in Chapter 4—there is no particular reason to assume that there is a clear distinction here at the level of *standards of evaluation*.<sup>75</sup>

In so far as we want to allow that states of affairs within a state *can* be more or less just, there might be a “truth” about the matter, in a minimal sense: we do not have to assume a Platonic version of moral realism in order to make such judgments. “Truth” in the sense used here can be intersubjective, expressive, locally bounded, or context-dependent; this still allows for “truth” in the sense required. Unless we want to deny the truth-aptness of moral statements outright and adopt either a radical non-cognitivist position or a value pluralist one with respect to morality, we can accommodate the idea of “better” and “worse” in the context of collective moral rules.<sup>76</sup> So, unless we are committing to one of those meta-ethical positions, we should assume that there is

---

<sup>75</sup> For a discussion of the fuzzy distinction between facts and values, see Hilary Putnam, *The Collapse of the Fact/Value Dichotomy: and Other Essays* (Cambridge, MA: Harvard University Press, 2002).

<sup>76</sup> I will return to this discussion in Chapter 4.

something to be discovered, and the reasonable disagreement we observe is really an epistemological problem.

Of course, this is not sufficient to rule out non-cognitivism or radical pluralism. However, we might not need to do that: in the next chapter I will pursue this argument further, but let me just mention for now that we might question the continuity of political decision-making with moral philosophy in the first place. Arguably, the standard of good decision-making in the domain of politics—especially decision-making under conditions of disagreement and uncertainty—is not the same as the meta-ethical standard by which we assess moral theories: *There may be better and worse political choices, independently of the truth-status of the moral theories underlying these choices.*

People may reasonably and deeply disagree about the final ends of politics, about ideology, or about the moral principles underlying political choice, but such disagreement leaves open the possibility that there might be some quite substantial agreement on *what is to be done*, politically speaking. In Sen's terminology, there might be disagreement at the *transcendental* level about what is just while there being quite some partial agreement at the *comparative* level.<sup>77</sup> This seems to me an extremely plausible way to look at things. We might not be able to agree on whether we should aim for a Rawlsian or a Nozickian distribution of goods, whether we ought to base our system of laws on the ten commandments or on Shari'a, but we might well be able to agree that a society in which there is little murder, theft, and few or no public health emergencies etc., is preferable to one where all these things are prevalent.

---

<sup>77</sup> Amartya Sen, *The Idea of Justice* (Cambridge, MA: Harvard University Press, 2009).



Disagreement at the transcendental level, even if it should be absolutely irresolvable on rational grounds (which could be taken to suggest some epistemic pluralism of values), does not entail that there is no objective truth about *what should be done*. Take for instance one of the most divisive political issues in the contemporary United States: the right to have an abortion. This seems like one of those quintessential issues where disagreement is deep and irresolvable, and it seems strange to argue from “rightness” in this circumstance, where there is a deep divide precisely about this “rightness” with respect to abortion. One side argues abortions ought to be available since there is a basic right of self-ownership of one’s body, the other side argues that abortions are absolutely morally prohibited and ought to be prohibited legally as well. The one side wants a state of affairs where abortions are legal and available, the other side wants one where abortions are illegal and unavailable. And both positions are based on deep moral convictions, which makes it seem that there is no standard on the basis of which we can come down on one side rather than the other.

However, this perspective is precisely the problem. Why should we measure the quality of a political decision by whether it manages to resolve *this* level of disagreement? The focus on the end-states desired by these two moral positions obscures a large area of potential *agreement*: presumably no-one thinks an abortion is in general a good thing; if not for moral or religious reasons, then because of the potential psychological and physical dangers of such a procedure. Therefore, there should be a general agreement that the *occurrence of preventable abortions* ought to be minimized: let’s say by free or inexpensive distribution of contraceptives, and perhaps by enhancing the public knowledge about it. While we might not be able to say which way the fundamental disagreement ought to be

resolved, there can still be a “right thing to do” here, which might well command agreement.<sup>78</sup> In Rawls’s formulation, the “overlapping consensus” between disagreeing parties might well be larger than the initial disagreement makes it look. David Wiens develops a similar idea with his “failure analysis approach” to evaluating policies and institutional regimes.<sup>79</sup> This means that not only does disagreement not entail that there is no truth about the matter, it also does not entail that there is no right thing to do given the disagreement. When faced with a political problem as well as disagreement about the final ends of politics, this does not mean that any action is as good as any other—and in particular it does not mean that inaction is as good as attempting to solve the problem as far as it goes.<sup>80</sup>

Finally, a related issue is the question of the *intransience* of disagreement; connected to this issue is the question what we are to make of the idea of ethical *progress* over time. On many moral issues, such as the rejection of slavery, pederasty, women’s equality, or religious tolerance, we tend to think that our beliefs are a distinct *progress* over what used to be believed; and beyond that, it even seems like nowadays there is near-universal agreement (in Western societies) on many of these issues. Yet at some point these issues were decidedly controversial—there was deep and deeply consequential disagreement as to the morality of slave-holding, as was about what should be done with people who don’t accept (or do accept) the doctrine of transubstantiation.

---

<sup>78</sup> The importance of “reframing” to overcome these situations of conflict is also discussed in detail in Amanda Roth, “Ethical Progress as Problem-Resolving,” *Journal of Political Philosophy* 20.4 (2012): 384–406

<sup>79</sup> David Wiens, “Prescribing Institutions Without Ideal Theory,” *Journal of Political Philosophy* 20.1 (2012): 45–70.

<sup>80</sup> I will pursue this general idea further and in more detail in the next chapter.

If we admit that *ex ante* disagreement about political issues may disappear *ex post*; i.e. if we allow that people might change their minds about a policy once they have experienced it, once more information has become available, the context has changed, or the general moral “knowledge” has evolved, the issue becomes much less of a problem. The fact that a divisive issue cannot be resolved *ex ante* does not mean that it is irresolvable: and it does not mean that none of the decisions can be *right* in the context of disagreement.

Consider the question of the official recognition of same-sex marriage: another extremely controversial issue. One of the key arguments of the opponents is that the practice of marrying someone of the same sex devalues the moral and symbolic significance of heterosexual marriages as well. Now, in most societies the legal recognition of same-sex marriages has not had that effect. Therefore, there is a case to be made that we should just try it out: since there is good reason to expect that no such devaluation effect will take place in practice, the key bone of contention might disappear after the policy has been tried. Assuming this is how it should turn out, there would be *ex ante* disagreement, but no *ex post* disagreement. If, however, the devaluation effect does occur, the issue is resolved the other way: the policy of instituting homosexual marriage should be reversed. Either way, the disagreement may disappear.

Another example: denying women access to higher education used to be justified on the basis of the belief that women’s emotional nature, temperament and thinking style meant they could not benefit from higher education and were better off with only limited education. But of course once women were able to access higher education, this turned out not true. After more than a century of women’s higher education we know that this

claim is patently wrong.<sup>81</sup> This particular *factual* disagreement has been largely resolved in Western countries.<sup>82</sup>

So people might, and in fact do, change their minds. Beyond this, however, *societies* change their (collective) minds as well even if no individual is inclined to do so. This is of course a slow process—as it often works by cohort replacement across several generations: but what is a disagreement today might not be a disagreement tomorrow.

If a political choice commands ex post agreement (if most people agree, “that was the right thing to do”), we should be able to say that it was the right choice despite the ex ante disagreement. And this rightness applies regardless of the method by which the question was decided procedurally. Now the intrinsic proceduralist might still think that the choice could have been *unjustified*, even though it turned out to be right. But in that case she would have to make a *positive* argument for the exclusive relevance of procedural aspects for legitimacy—the difficulties with such an argument have been addressed above. *Pace* Waldron, disagreement does not *automatically* entail that epistemic justifications are self-defeating.

What this brief discussion shows is that *disagreement*, which manifests itself as *uncertainty* as I have spelled out, does not imply that implementing a policy that comes down on one side of an issue is *necessarily* unjustifiable in instrumental terms and therefore in need of some other source of legitimacy. Despite the fierce opposition, mandatory smallpox vaccination, as it turned out, would have been the right thing to do whether or

---

<sup>81</sup> If anything, evidence suggests that the reverse is probably true.

<sup>82</sup> Sadly, of course this does not mean that there are no opponents to women’s access to higher education anymore—however, at least they have to think of a different gratuitous argument to support that view.

not it was decided democratically; arguably, in terms of lives saved, the eradication of smallpox was one of the best policy decisions of all time.<sup>83</sup> The difficult bit, of course, is *how* to give such a justification: I address this in the next chapter.

This discussion of disagreement has challenged the thesis that deep disagreement entails that epistemic justifications of political legitimacy are self-defeating. It has done so on three counts: firstly, I have argued that *reasonable* disagreement does not rule out that there may be a truth about the matter, and that we should see disagreement as an epistemological problem, namely a problem of *uncertainty*. Secondly, I have argued that disagreement about *final ends*—what I have called “transcendental disagreement” following Sen—does not automatically imply the same level (or depth) of disagreement at the practical-political level about *what is to be done*. And thirdly, I have argued that ex ante disagreement about what is to be done need not imply that any decision is *prima facie* unjustified unless procedurally legitimated.

## 10. Sources of Disagreement

To recapitulate the whole argument so far: (1) Theories that try and justify political legitimacy on procedural grounds face a dilemma: if they are based on purely formal principles they are indeterminate and hence insufficient to establish legitimate authority, but if they are based on more substantive values, epistemic considerations must come into play. (2) Hybrid theories that accept a lexically secondary role for epistemic

---

<sup>83</sup> see for instance Stefan Riedel, “Edward Jenner and the History of Smallpox and Vaccination,” *Baylor University Medical Center Proceedings* 18.1 (2005): 21–25.

considerations do not work; they either collapse into a pure procedural theory; or they have to accept that epistemic considerations may in principle trump procedural ones; given that argument (1) has illustrated the difficulties with pure proceduralism, we can see therefore that epistemic considerations are of crucial importance when evaluating political procedures. Finally, (3) reasonable disagreement does not necessarily imply that any epistemic justification is self-defeating; epistemic considerations are possible, even though they have to take into account the fact of disagreement and pluralism.

In sum, therefore, I have tried to argue that democratic legitimacy, if democracy can be legitimate at all, has to be based fundamentally on its epistemic capacities, its capacity and tendency to produce good outcomes. In other words, when we try to figure out whether a democratic reform is a good idea or not, when we try to design the best democratic institutions, it is important to keep their *epistemic* capacities, their information-processing and decision-making abilities, in mind; not only that, the epistemic capacity is a *central concern* we should have when doing that.

What this chapter so far has *not* established is *how* one can make a convincing argument of this kind, and what it would look like (and which institutional setup it would recommend). I will pick up this question in chapter 3. As we will see, this is difficult, because of the issues that have come up throughout the past chapter, and because political problems are very complex things indeed. This last section therefore explores the difficulty of the task of the would-be epistemic democrat.

So, if you buy the argument so far, it seems that the epistemic reliability of a political system is of critical importance for its normative evaluation. However, there is a

crucial difficulty: we don't know what reliability consists in. We often cannot say *ex ante* what is the right thing to do: there is extensive and deep disagreement about this. And even given this disagreement, as I have shown, it is likely that there are better and worse things to do; however, we cannot be sure what they are. We face *epistemic uncertainty* about what to do—even though many people of course think they know exactly what should be done (or at least claim the same). This means, or so it seems, that we cannot evaluate different procedures as to whether they tend to do this right thing or not.

But what is the nature of this uncertainty? Why is it that often we do not know what to do in the political realm? There are a number of possible answers to this question: and importantly, *how we think about this uncertainty determines (to quite a large extent) how we think about epistemic reliability*. What are political problems (and their solutions) *like*? To answer this question, I'll take a short detour into complexity theory. In particular I submit that the uncertainty here is a product of what Elster calls a “double indeterminacy”—an uncertainty about the validity of normative views and an equal uncertainty about the truth of factual claims.<sup>84</sup> There are three potential explanations of such uncertainty. The first two are relatively unproblematic from the standpoint of democratic theory; however, since the uncertainty we encounter in the political realm frequently is of the third sort, this is a serious issue for democratic theorists, and therefore also for the theory of political legitimacy advanced in this dissertation so far.

Firstly, we might think that uncertainty (and therefore disagreement) is just due to honest mistake or involuntary ignorance. Many people have opposed views about what should be done, and this could be because we have made failures in reasoning, have false

---

<sup>84</sup>Jon Elster, *Securities against Misrule* (Cambridge: Cambridge University Press, 2013). Introduction

factual beliefs or miss important information that would remove the uncertainty. Therefore, what seems like an irresolvable clash of opinion or an unsolvable problem could presumably be resolved by correcting the mistake. For instance, assume we don't know whether instituting compulsory vaccination is a good thing or not. Now assume my opposition to compulsory vaccination against smallpox might be based on a mistaken belief that vaccination is excessively harmful; which in turn might be based on a failure of reasoning, such that anecdotal evidence (“but I know someone who was harmed by vaccination”) is given more weight than statistical evidence. But in principle, this uncertainty could be resolved; we “just” need to straighten out our reasoning.

It is clear that if this kind of thing is the main source of our ignorance, what this calls for is more active deliberation within our political processes. Presumably, the give and take of arguments and evidence in deliberative contestation will gradually remove the uncertainty associated with unintentional mistakes—assuming that the deliberation is free of the grossest power inequalities, that people are focused on actual reflection about the issue at hand, and that we can get over the other potential pathologies of group deliberation.<sup>85</sup> This process might be approximated in representative assemblies, or mini-publics to the extent that they really do argue and not bargain.<sup>86</sup>

A second view of this uncertainty would be that it is due not only to involuntary mistake but to inherent cognitive limitations of humans. For one thing, there are of course many very difficult problems, and as already mentioned above I might not have the time

---

<sup>85</sup> E.g. groupthink, information cascades, polarization. See for instance, Cass R Sunstein, *Going to Extremes : How Like Minds Unite and Divide* (Oxford: Oxford University Press, 2009).

<sup>86</sup> On the distinction, see Jon Elster, “The Market and the Forum,” in *Foundations of Social Choice Theory*, ed. Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986), 103–132.



or necessary expertise to delve deep into taxation law, healthcare, or banking regulation. My view will therefore be partial almost by necessity. But even for notionally “simple” questions, there are some identifiable issues with our cognition. These limitations of course tend to manifest as “biases,” and have been documented in many important studies.<sup>87</sup> Prospect theory, among many other things, of course famously shows that people are subject to framing or anchoring effects, which may lead to inconsistency in opinion.<sup>88</sup> Furthermore, several scholars have written about the human tendency to find patterns in essentially random (unpatterned) events, or impose explanatory narratives on the chain of events that rationalize a certain development by fitting it into that narrative.<sup>89</sup> Accordingly, we tend to assume that the pattern we erroneously impose on the past will continue in the future, misleading us into making false predictions. Thus, for example, I might put down my personal health not to statistical luck but to my lifestyle; I might tend to construct a narrative around how my decisions to exercise, etc., *caused* my current health. On the basis of this narrative I then oppose universal healthcare since I believe that everyone’s personal decisions, not objective risk, are to blame for bad health. Another example: assume I am wealthy; then I might attribute my wealth exclusively to my great decisions and hard work in the past; thus ignoring the many other people who made similar decisions and worked even harder but were not in the right place at the

---

<sup>87</sup> Daniel Kahneman and Gary Klein, “Conditions for Intuitive Expertise: A Failure to Disagree,” *American Psychologist* 64.6 (2009): 515–526.

<sup>88</sup> For instance, one and the same outcome can be evaluated differently, depending on how it is described. As Kahneman shows, many people regard an action that will be “saving 400 of 600 people” as better than one that will be “losing 200 of 600 people.”

<sup>89</sup> Duncan J Watts, *Everything Is Obvious: Once You Know the Answer*, 1st ed. (New York: Crown Business, 2011); Nassim Nicholas Taleb, *The Black Swan : the Impact of the Highly Improbable*, 1st ed. (New York: Random House, 2007).

right time. Accordingly, I might form the opinion that people who are poor only have themselves to blame, and that inequality is very well justified.

In those cases, the uncertainty results from individual cognitive limitations; from the fact that no-one can easily (or perhaps not at all) take the objective perspective on entire political problems, let alone politics as a whole. Political decisions have to be made for all, while our evaluations to some extent subjective. If cognitive limitations of this sort are the main source of uncertainty, it is again mainly a failure in proper reasoning that leads to the uncertainty. However, while real deliberation might ameliorate some of the reasoning failures due to cognitive limitations, it also might not. Expertise—especially the capacity to examine and evaluate evidence—might have to play a much larger role in a solution to this particular problem of uncertainty. We might need, as it put nicely by Kitcher, a *cognitive division of labor*.<sup>90</sup> The key is to reach the right *social* knowledge:

“Our most recent observations concerning the vast heterogeneity of knowledge required for governance make it clear that epistemic adequacy will hinge, not on mastery of any one body of knowledge, but rather on our ability to coordinate many discrete bodies of knowledge, diversely distributed across the political community. It becomes clear, in other words, that the proper epistemic perspective can only be social.”<sup>91</sup>

This does not mean that we should abandon democratic structures in favor of some form of the dreaded expertocracy, but it might call for some degree of professionalization of politics; and it may support a trustee model of representative politics over a mandate model, let alone more direct democracy. In certain extreme cases, however, this may call for a walling off of certain policy areas from democratic control altogether—as has happened for instance with monetary policy in most countries by now. From this

---

<sup>90</sup> Philip Kitcher, “Division of Cognitive Labor,” *The Journal of Philosophy* 87.1 (1990): 5–22.

<sup>91</sup> Fuerstein, “Epistemic Democracy and the Social Character of Knowledge.” 78

perspective, the role of media influence in politics also become a key issue; the more important framing becomes relative to substance, the more important the framing agents become relative to content.

Finally, however, the answers to political questions may just simply be *uncertain* in themselves. Political questions are difficult, the behaviour of masses of people and complex systems like entire economies are all but unpredictable, and the consequences of purposive political action may be unforeseeable.<sup>92</sup> Since the consequences of policy depend fundamentally on the people's reaction to it, and since people's behavior is influenced by a constellation of many different factors, whether or not a policy will reach its stated goal also depends on precisely what constellation of factors obtains.

In short, the success of policy depends on the behavior of the society/economy/ecosystem within which it is implemented.<sup>93</sup> These are *complex adaptive systems*: and the behavior of complex adaptive systems is particularly difficult to predict. A complex system is characterized by a *diversity* of types within the system that have a large number of interconnections; such a system is *adaptive* if the units within it adapt their behavior to perceived circumstances. We might say: the people within a system differ fundamentally in their preferences and beliefs—that is, their expected behavior; and beyond that they adapt their behavior to the perceived and expected behavior of other agents within the system. Minute changes can therefore have extreme effects, and large changes may have little effect, it all depends how the system adapts.

---

<sup>92</sup> Robert K Merton, "The Unanticipated Consequences of Purposive Social Action," *American Sociological Review* 1.6 (1936): 894–904.

<sup>93</sup> A more detailed discussion of this follows in chapter 4 below.

The problems arising from complexity and difficulty in politics have recently become known as the issue of “wicked problems,” that is, problems that exhibit difficulty on a number of different interacting dimensions.<sup>94</sup> The key insight for our purposes here is that *if* problems are complex and the subjects of policy are complex adaptive systems, then *solutions* will require entirely different strategies and institutional mechanisms than when the difficulty of the problem is “merely” due to error in reasoning or cognitive limitations. As Keynes puts the point:

“The expectation of life is only slightly uncertain. Even the weather is only moderately uncertain. The sense in which we are using the term [uncertainty] is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth owners in the social system, in 1970. About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know. Nevertheless, the necessity for action and for decision compels us as practical men to do our best to overlook this awkward fact.”<sup>95</sup>

In other words, the policy-maker faced with such problems will have to act *despite* the uncertainty, or, to rather *act taking uncertainty into account*: The uncertainty will not go away. But it is not immediately obvious which institutional mechanisms would be useful for that purpose—indeed it is not even certain what it would mean to act well, *overlooking the uncertainty*.

---

<sup>94</sup> Horst W J Rittel and Melvin M Webber, “Dilemmas in a General Theory of Planning,” *Policy Sciences* 4.2 (1973): 155–169.

<sup>95</sup> John Maynard Keynes, “The General Theory of Employment,” *The Quarterly Journal of Economics* 51.2 (1937): 209–223; cited in Sandra D Mitchell and Wolfgang Streeck, “Complex, Historical, Self-reflexive: Expect the Unexpected!,” *MPIfG Working Paper* (Cologne: Max-Planck-Institute for the Study of Societies, 2009). 7

## 11. Is Politics Really That Difficult?

This last section was concerned with establishing that instrumental justification theories do have to take deep uncertainty into account. This is predicated, of course, on the assumption I have outlined, that political problems really are that difficult and unpredictable that we don't know what to do. Now one might object that I overstate the problem; or at least exaggerate the universality of the issue.

For that reason I will consider three possible objections. Firstly, one may think that most, or the most important, political decisions are not actually that complex. This view would argue that complexity and uncertainty are issues that bedevil philosophy of science, and perhaps sciences more generally, but that the questions with which politics is concerned are not like those scientific problems. We might think that politics is concerned mostly with, or can meaningfully be reduced to, simple discrete (often local, often less consequential) decisions; specifically with balancing preferences, not with estimates of how complex systems behave, and that the level of accuracy demanded by me is ludicrously high. Politics is frequently concerned with whether to build a football stadium or a swimming pool, whether to fund the National Endowment of the Arts or the National Institute of Health, etc.—decisions whose consequences are neither radically unpredictable nor of a magnitude that has potentially catastrophic consequences.

I agree that politics is concerned with simple questions as well as complex ones, but I would stress that the latter *are* part of politics as well. Emergency preparation, environmental protection, macro-economic policy, long-term fiscal policy and population/immigration control are all fields that are essential to managing a complex

system such as a modern society. It is possible that these questions make up a small minority of political problems (although I doubt it). I do want to argue though that the reliability of a political decision structure should be evaluated mostly by its ability to acceptably deal with these difficult problems, not largely how it deals with the easy or less consequential ones. Easy questions, almost by definition, will not allow us to differentiate epistemic mechanisms according to their advantages. For easy problems, any decision mechanism is as good as any other—or rather, in such cases selection criteria other than epistemic advantage may come into play more strongly: when we debate the placement of the football stadium it might be most important to get all stakeholders involved, more than getting the placement “objectively” right—or more precisely, stakeholder involvement, is so far as it is necessary for the success of a policy, *is* a part of what it means to be objectively right in such circumstances.

A related objection might be that political decisions are in any case concerned less with causal truths, but with moral ones: politics should be concerned with just decisions, not with factually accurate ones. In this view, an adequate epistemic procedure is one that tends to discover normative truths—and it is this latter we should care about as instrumental democrats. And, so one might argue, normative questions do not suffer from the same kind of unpredictability that I have identified as a problem. They might be difficult, but they are not *complex* in the same way. With respect to this I would reply that the task of answering normative questions in politics only *looks* deceptively simple. For one thing, identifying what is just or right is only the first step: the second is the question how to *realize* justice or the correct set of rights in society. This, as it were, depends again on the behavior of complex adaptive systems that is at the root of our uncertainty. Consider

once again Estlund’s notion that one of the main tasks of politics is the prevention of *primary bads*. They include famine, genocide, political collapse, economic collapse, or civil war. Now while it is true that the question whether famine is good or bad is a simple one—one that everyone can figure out for themselves and where everyone is likely to agree. However, this question is not the most important one. What we really want to know is *how* to prevent famine: and if anything, this seems like an extremely complex question.

Additionally, of course, there are many *pro tanto* political values that are potentially in conflict. As G.A. Cohen put it, the fact that justice requires A does not immediately imply that we should do A.<sup>96</sup> There are other things we value, freedom, privacy, intimacy, love. Justice does not automatically and at all costs override all other values, and the question *what is to be done* is not resolved once we figure out the “simple” question of what justice requires.

A third objection may be that politics is actually *responsible* for much of the uncertainty; and consequently, political decisions can be used to *reduce* that complexity as well. For instance, the complexity of the road system is a product of the rules of the road; and the reason that markets are so difficult to predict is of course a consequence of the *laissez-faire* approach to economic policy in most liberal democracies. Thus, we might think that those complicated problems actually have potentially simple (and probably technological) solutions.<sup>97</sup> However, as Scott Page points out, when we are faced with a complex adaptive system, the complexity does not disappear even though we might try

---

<sup>96</sup> see G A Cohen, *Rescuing Justice and Equality* (Cambridge, MA: Harvard University Press, 2009).

<sup>97</sup> On the problems with this view, see Evgeny Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism* (New York: PublicAffairs, 2013).

and implement radically simplifying policies. To give Page's example: we could easily radically reduce the complexity of the traffic system by allowing only high-occupancy-vehicles (HOVs) on the road. There would be many fewer cars on the road, and the task of going from A to B would be much easier. However, if I have to have at least 3 passengers to be allowed on the road, the task of *planning car journeys* becomes much more difficult for the individual.<sup>98</sup> When lots of people want to get to many different places via a limited number of roads, the resulting system will be complex whichever set of rules we impose; the question is just where the complexity will manifest itself.

We might judge a problem to be relatively simple, and it can turn out to be fiendishly complex—like the task of the Soviet central planning agencies of balancing the inputs and outputs of the different firms, manufacturers and consumers in the economic system. In a planned economy this seems straightforward enough: after all there is a control over which goods are produced and therefore a limit to the diversity of the economic system. However, this transfers the complexity of the coordination task from individual firms to a central planning system. Even for the limited number of different goods available in the Soviet system, this task turned out so complex that the Soviet economist Abel Aganbegyan said in 1964, that by 1980 the whole population would have to be employed full-time in finding the general equilibrium for the plan.<sup>99</sup> Simple solutions, frequently do not resolve the uncertainty, they just displace it.

---

<sup>98</sup> Scott E Page, "Uncertainty, Difficulty, and Complexity," *Journal of Theoretical Politics* 20.2 (2008): 115–149.

<sup>99</sup> quoted in Francis Spufford, *Red Plenty* (London: Graywolf Press, 2012) Note to p. 219.



## 12. Conclusion

This leaves us in quite a bind: we think that a political procedure's epistemic reliability is essentially important for its justification. But then, if problems are complex, we cannot say what the solutions to our political problems are; and therefore it seems difficult to say which procedures would *tend* to find the right solutions.

To give up on instrumental justification, as I have tried to argue so far, is unattractive: pure proceduralist justifications are too indeterminate to be of much use in looking at existing political procedures or coming up with possible alternatives. And there is nothing in formal procedural values themselves that would suggest one interpretation rather than another. There are many kinds of fair procedure, and the value of "fairness" itself cannot tell us which one is applicable for a given context. Procedural justifications therefore need to refer to the *purpose* of the procedure, its *instrumental value*; and I have argued that this should be its *epistemic reliability*, its capacity to find the *right* solutions to political problems.

I have also argued against hybrid theories that would assert a procedural value and the use of an epistemic criterion to choose from the range of procedures allowed by the initial criterion. The key reason is that the basis for asserting the lexical priority of the procedural criterion also defeats the secondary epistemic criterion. Therefore, I must either allow the indeterminacy of the pure procedural justification, or allow that epistemic considerations, at least in principle, may defeat procedural ones.

Finally, I have argued against procedural justifications as a *residual value*. The facts of disagreement, pluralism and uncertainty about the truth do not imply that we cannot

assume that there would be a *right thing to do*. We do have to take these “circumstances of politics,” as Waldron calls them, seriously, but this just means that insofar as there are better and worse decisions in politics, they are right *given* disagreement and uncertainty.

But, as the last section has shown, while disagreement and uncertainty do not entail that there is no right thing to do, they sure make it difficult to see what that might be. Especially if we see political issues as *inherently* difficult and beset by complexity and unpredictability, we are faced with the problem that we don’t know what the right thing to do will turn out to be. Accordingly, it is difficult to define epistemic reliability for a given procedure at all.

It looks like the only alternative would be *philosophical anarchism*, the view that no government can be justified and therefore no government is normatively more desirable than any other (objectively speaking); but this view seems counter-intuitive; the concept of political legitimacy is of essential importance in actual political debate, perhaps we are not quite ready to dismiss it as so much cheap talk. Especially democratic politics seems, in many ways, *more* legitimate than other forms of government.

The only way out, therefore, is to try and assert democracy’s *universal epistemic reliability*; that is, democracy’s reliability for *all* possible (or at least for all likely) scenarios of what the right solution may turn out to be. We have to make an epistemic argument without presupposing any *particular* truth or conditions of truth. What is epistemic reliability, for political contexts? How we can make such an argument, and under which conditions, is the subject of the next chapter.

---

## Robust Universal Reliability of Democracy

### 1. Is “Universal Reliability” Possible?

The last chapter established that outcome considerations play an important moral role in the legitimation of political authority, and this is *despite* the problem of pervasive disagreement and uncertainty about the quality of outcomes. However, we still do not know what precisely this means. The problem already mentioned at the end of the last chapter was this: if we want to justify democracy on an instrumental basis (by the outcomes it produces), we need to claim and argue that democratic mechanisms have *robust reliability*, that is, they have to make good decisions under all circumstances we could reasonably expect—this also means that democracy should be expected to get good outcomes, relatively speaking, whatever “good” will turn out to mean.

This is a formidable task, and a high standard by which to judge democracy’s legitimacy. However, recall that the criterion of legitimacy I am using—robust instrumentalism—is a variant of justificatory liberalism: for a decision procedure to count as legitimate it must be in principle acceptable from all reasonable points of view.

Therefore, in order to construct a convincing instrumental justification of democracy, the claim to democracy's reliability itself must be *robust*, that is in principle acceptable from all reasonable points of view.<sup>100</sup>

*The demands of justification*

Now, the last chapter has shown that a justification that is acceptable in this sense cannot refer *only* to procedural features of the democratic system that is to be justified. There must be some argument that the proposed democratic system would tend to make good or at least acceptable decisions. Since, however, there is uncertainty and disagreement about what a good decision is, this must be what I call a "universal reliability argument." Democracy must have an "edge" with respect to producing good decisions, *whatever that may turn out to mean* in the end. In other words, any argument for the epistemic edge cannot be predicated on some *specific* definition of what is to be discovered by the procedure.

This is because it seems that we cannot form the expectation that a particular political procedure or form of rule will (tend to) produce the right outcomes, unless we already know what those outcomes are, and how they could be produced. But the precise problem was that there is disagreement and uncertainty about this. As Waldron very aptly recognizes: if a procedure is intended to resolve such disagreement, the justification of the same procedure cannot be predicated on an *ex ante* assertion of the right answer.<sup>101</sup> Having the right answer would eliminate the need for such a procedure. It seems

---

<sup>100</sup> Gaus, *Justificatory Liberalism : An Essay on Epistemology and Political Theory*; Estlund, *Democratic Authority: a Philosophical Framework*.

<sup>101</sup> See also the quotation above from Waldron, *Law and Disagreement* 253.

problematic, say, to justify the imposition of a particular authoritative procedure in spite of disagreement by pointing to the fact that it resolves that disagreement in a particular way. If we look at it this way, an epistemic *justification* of democracy, given the fact of pluralism and uncertainty, seems to be doomed from the start: why would you accept my justification of democratic authority on the basis that it will reliably discover political truth T, unless you already accept the truth of T? And if you (reasonably) have a different opinion about T, why would you accept my justification?

A different way of putting this is that we need to define the standard by which we judge the “reliability” of democracy, namely the tendency to reach good decisions in a way that is consistent with the uncertainty and disagreement we observe in political matters. Using Gaus’ expression, this is the question of the “epistemic test” and how it is to be defined.<sup>102</sup> The definition of the epistemic standard is indeed a crucial normative step, since the whole working of the argument turns on it. We could of course define epistemic advantage in a way that democracy automatically comes out on top: if “good decisions” means “doing what the majority vote determines” then majority rule is by definition best at this task. At the other extreme, we can also define a “a good decision” in a Platonic way as “deciding in accordance with the Form of the Good,” where only true philosophers have access to this Form: this would of course justify a system of rule by true philosophers who have this access.

Such “bespoke” definitions violate the standards of robust instrumentalism. Restrictions on how the “reliability” can be defined are given by the features of what we

---

<sup>102</sup> Gerald Gaus, “On Seeking the Truth (Whatever That Is) Through Democracy: Estlund's Case for the Qualified Epistemic Claim,” *Ethics* 121.2 (2011): 270–300.

think is a valid epistemic justification of democracy in the first place. As per the assumptions of justificatory liberalism and reasonable disagreement, the concept of epistemic reliability has to be defined in a way that the resulting justificatory argument is consistent with these commitments. Could a reasonable person accept democratic political authority over her, on the basis of such an argument? For current purposes I shall identify three minimal desiderata of a successful justification of this kind. It seems to me that any reasonable liberal point of view would at least demand these from any justification they could accept. These are *rationality*, *optimality*, and *robustness*. Any valid argument grounding the legitimacy of democratic authority must, at a minimum pass these tests.

By rationality I refer to the idea that in order to fulfill the condition of legitimacy, democracy must enjoy an advantage over alternative systems with respect to the *ends* of politics, or what we may call *normative* commitments, as well as, if necessary, *causal theories about the world*.<sup>103</sup> An epistemic argument is not complete if we argue that our epistemic procedure is good at identifying the right goals of policy (i.e. that it discovers the correct normative beliefs), without saying anything about its efficiency, as a system, to realize those goals (i.e. its ability to base policy decisions on bring about these goals). If democracy always exhibits good intentions, but systematically gets it wrong when following on those intentions, a successful justification cannot be based on that. Conversely, a system that exhibits great instrumental rationality, i.e. has appropriate

---

<sup>103</sup> I adopt the term “rationality” for this requirement from List and Pettit, *Group Agency : the Possibility, Design, and Status of Corporate Agents*, 67.

causal beliefs about the world, cannot be considered justified unless the rationality is directed towards the right goals.

Optimality refers to the simple idea that instrumental reliability has to be defined such that democracy is the decision method that is *best* at getting things right (either among all possible institutional setups, or relative to the relevant reference class), or that it alone passes the threshold. A justificatory argument ought of course successfully demonstrate that the political regime to be justified actually passes the bar set by its own principle of legitimacy. Hence, a successful outcome-based argument for democracy will have to show that democracy actually is better than its feasible alternatives at producing good decisions or outcomes. A weaker form of this claim, *weak optimality*, might be that democracy is *at least as good* as the best feasible alternatives, which would allow for ties among the best regime forms (e.g. democracy and alternative procedure P are “joint top”).

Finally, this claim of democracy’s optimality must be *robust*. This follows from the demand that political authority be justifiable from all reasonable points of view. This is a slightly more complicated concept. It reflects the moral importance of the high degree of uncertainty about the truth, as well as the fact of reasonable disagreement in current societies. More specifically, robustness as it is used here means that for a given standard, the case that whether democracy is indeed uniquely optimal according to that standard should not depend on many contentious assumptions<sup>104</sup> about people’s capacities and the structure or difficulty of political problems—in short, about what the political truth actually is. There are two aspects to this. First, in social choice theory a collective decision

---

<sup>104</sup> That is, assumptions that could be reasonably rejected.

mechanism is said to be “robust” when it produces satisfactory outcomes for any possible profile of inputs across the population.<sup>105</sup> Here I want to use robustness in a weaker sense as the requirement that a social choice mechanism will produce epistemically desirable outputs, or “good decisions,” for any profile of democratic inputs (e.g. sets of preferences) we might reasonably expect to obtain within a population. We are not sure how competence, knowledge or motivations are distributed across the population, and our justification must “work” under a wide range of these. For instance, if democracy only produces good outcomes whenever people are especially wise—and we have no independent reason to expect that people are so wise—then it is not *robust*.

The second aspect of robustness is what has already been mentioned above: that a given justification may not be predicated on any particular definition of what a “good decision” is. One reason is Waldron’s worry above: that an argument for a procedure that is supposed to resolve a disagreement cannot itself depend on the correctness of one of the sides of the disagreement. Beyond that, in purely logical terms, of course, any such claim is not only not robust, it also violates the optimality condition: if I assert a substantive “right” solution—then presumably I already have a reliable method by which I know what is right. Otherwise, why should anyone place trust in that assertion? And if I already have a highly reliable procedure, how can I claim *optimal* reliability for democratic procedures?

Why should a successful justification be robust in this sense? Essentially, because we do not know, *ex ante*, who has the right answer: as J. S. Mill puts it, it may be equally likely “that the received opinion may be false, and some other opinion, consequently,

---

<sup>105</sup> *ibid.*, 67-70.



true; or that, the received opinion being true, a conflict with the opposite error is essential to a clear apprehension and deep feeling of its truth.” Or, it may be the case that “conflicting doctrines, instead of being one true and the other false, share the truth between them; and the nonconforming opinion is needed to supply the remainder of the truth, of which the received doctrine embodies only a part.”<sup>106</sup> The point is, we don’t know which of these conditions hold, therefore an epistemic procedure can be justified only when it is successful under either condition.

How strictly we interpret the robustness requirement of robust instrumentalism depends on how “ideal” or “aspirational” we think our theory of justification should be. Should it be an ideal to which to aspire, or should it offer concrete improvements in the current situation? Perhaps we think we are free to assume that people will generally be rational and not overly malicious in their behavior, and it is acceptable that democracy should produce reliable results only under these conditions. However, an argument becomes questionable when our reliability argument depends on assumptions how individuals have certain cognitive capacities (or motivations to improve their cognitive capacity) that they are unlikely to actually have—even under ideal conditions. To put it in the words of *Federalist* 51: if men were (epistemic) angels, no government would be necessary. And if the acceptance of a justification of democracy is predicated on the acceptance of such assumptions, it might be entirely reasonable to reject it.<sup>107</sup> In order to

---

<sup>106</sup> John Stuart Mill, “On Liberty,” in *Utilitarianism; On Liberty; Considerations on Representative Government; Remarks on Bentham’s Philosophy*, ed. Geraint Williams (London: J.M. Dent, 1993) 113.

<sup>107</sup> The issue is that such a form of justification must balance on a fine line: one the one hand, if we assume behavior that is too ideal, democracy’s quality will no longer be distinguishable from that of other potential systems: if people were epistemic angels, any form of government (or none at all) will probably perform equally well, and there is no particular reason why democracy should serve as the aspirational ideal, rather than, say, benign expertocracy. On the other hand, if we allow just enough cognitive limitation in our

avoid such issues, robustness requires that the mechanism to be justified produce the desired epistemic effects under all *reasonable* scenarios.

*The reliability claim*

Now the task would be to find an instrumental justification of democracy based on a definition of “epistemic reliability” that is both convincing *and* satisfies the three criteria. So first, we should have to clarify whether political questions actually admit of better and worse solutions, and second, we have to find a way to flesh this out into a successful argument.

I have already discussed in section 9 of chapter 2 some reasons why it does not follow from the fact of disagreement that the concept of “better outcomes” is inapplicable to democratic politics. I have argued that disagreement ought to be seen as an indication that there is an *epistemic* problem, a problem of uncertainty about what good outcomes are and how to bring them about, not (necessarily) an indication that there are no better and worse outcomes. As it were, the empirical observation of disagreement and uncertainty is not equivalent to the normative judgment that there is no rational resolution to this situation, resolutions that admit of degrees of quality.

The point that situations of disagreement and uncertainty call for further inquiry seems obvious enough—almost trivial—when applied to other contexts: science, engineering, indeed most areas of our daily lives: disagreement is taken to call for further

---

model that democracy will do better than its alternatives, then there is the question why exactly that level rather than any other is deemed a reasonable assumption—and whether it is not just arbitrary.

(ideally somewhat intelligent) inquiry in order to resolve that disagreement, rather than a recognition that there is no objective answer to be had. Disagreement about, say, the nature of gravity, is not taken to indicate that there is no correct answer about this; and neither is it usually taken to imply the strong epistemological thesis that even if there was an answer, it cannot be discovered.

However, in normative contexts such as morals and (maybe) politics, this looks a little different. Where personal preferences and/or emotions are involved, it seems more difficult to conceive what a “right answer” or an “objectively good outcome” would look like. Much like there being no point to arguing about matters of taste, we might think there is no point to arguing about *goals*, of individual or collective life. For that reason, this chapter is also concerned with the question of whether there is a difference between disagreement about *ends* and disagreement about *facts*. The argument of this chapter will be that there is no such relevant difference—at least in the political context. Ends, goals, moral principles—in the context of politics—admit of *inquiry* just as facts or causal relationships.

This is not to say that this is easy, or even always feasible. Disagreement—understood as uncertainty—is a serious issue. This is precisely why it has to be addressed in a political context. When faced with uncertainty, the adequate response, as it were, is to take the rational decision *given that uncertainty*.

On this basis therefore I construct the *robust* outcome-based theory of democratic legitimacy we have been looking for in the second chapter. Remember the issue was that we value the quality of political outcomes very highly, but that there seems to be too

much uncertainty and disagreement about quality to come up with a definite theory of legitimacy purely on an outcome basis.

The next sections of this chapter are therefore concerned with establishing what we might call a *possibility result*: that it makes sense to speak of better and worse decisions, and hence of better and worse political procedures, *even if there should be disagreement and/or uncertainty about that*. First, I consider this question from within a purely social choice perspective that takes *preferences* as ultimate. Here I argue that even under these assumptions, in a limited way we can admit the *possibility* of better and worse decisions. Second, I consider whether we may come up with a more substantive conception of the quality of political outcomes, based on a conception of the “public interest” and its relationship to everyone’s individual interests. This argument leads me to the principle that *it is in everyone’s interest as a reasonable citizen that there should (a) exist a state that (b) is forced to (robustly) resolve common problems*. This principle becomes the basis for the *robust instrumental* justification of democratic authority.

On this basis, the chapter then goes on to define what this principle means in detail. The following sections establish that the appropriate response to facing situations of disagreement and uncertainty is *not* to reject any claims as to the quality of political outcomes, but to engage in further *inquiry* as to what better outcomes may be. The rest of the chapter then argues that this further inquiry should be *experimental*, and that on this basis we should formulate a *pragmatist* response to the problem of democratic legitimacy. I propose, therefore, a *pragmatic* version of an instrumental justification of democratic legitimacy.

This is a key point: such a justification is both appropriately responsive to the concern with the quality of political decision-making and the fact of uncertainty about what that quality actually consists in. The epistemic reliability of this argument is grounded not in democracy's inherent tendency to always make good decisions, but in its *experimental* and *adaptive* capacity over time to progressively resolve political problems.

The theory may be seen in the spirit of Tocqueville's argument about democracy: "If democracy has more opportunities for making mistakes, it also has a better chance of returning to the truth when the light dawns because, in general, it harbors no interests which oppose the majority or challenge reason. But a democracy cannot lay hold upon the truth except by experience and many nations might perish while they are waiting to discover their mistakes."<sup>108</sup> If democracy is to be justified instrumentally, and on the basis of the assumptions of justificatory liberalism, it has to be on the basis of the experimental methodology that democratic procedures represent.

The *pragmatic argument for experimental democracy* will be discussed in much more detail in later sections of this chapter, but here is the basic structure:

- (1) Only those political procedures are justified that can reasonably be expected to be robustly better than their alternatives at progressively overcoming political problems as they arise.
- (2) Given the extreme uncertainty that surrounds political problems—in the sense of both disagreement and ignorance about the problem structure and/or possible solutions—*experimental* methods of policy-making can reasonably be

---

<sup>108</sup> Alexis de Tocqueville, *Democracy in America*, ed. Isaac Kramnick, trans. Gerald Bevan (London: Penguin, 2003) 263.

expected to progressively overcome political problems, better than alternative systems.

- (3) Experimental methods of policy-making have certain functional requirements.
- (4) (Some) democratic political structures (those I will call elements of “experimental democracy”) are best suited to fulfill the functional requirements of an experimental model of policy-making.
- (5) Institutions that fulfill the functions of experimental democracy enjoy legitimate political authority.

This overview of the argument is fleshed out in much greater detail below. It needs to be defined what we mean by “political problem,” and especially also how we can “progressively overcome” them. Clarification is also in order with respect to what “experimental” methods of policy-making actually are—and whether are really justifiable. Finally, the connection between experimentation and democracy in the sense of (4) has to be made clear. Sometimes writers seem to assume this as a matter of course; that democracy is by its nature potentially an experimental and adaptive form of policy-making.<sup>109</sup> However, it seems obvious to me that not all systems that can be described as democratic are actually always working as expected. The question is what can be done to make democracy fulfill its potential.

For that reason, one advantage of conceiving of the normative function of democracy in this experimental way is that it gives us a new perspective to evaluate

---

<sup>109</sup> See for instance Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*; Anderson, “The Epistemology of Democracy.”

different existing democratic systems: those that better embody the experimental function required for legitimate authority and those who do worse. In chapter 6 I will suggest some institutional pointers we can use to perhaps make such judgments.

*The argument coming up*

In order to establish the pragmatic version of democratic justification, this chapter proceeds as follows: Section 2 of this chapter is concerned with the possibility of making rational choices under unresolved conflict. The argument is that it is *false* that there are no possible standards of better or worse under conditions of *prima facie* irresolvable disagreement. The individual case is an instructive analogue here: if there is an irresolvable conflict between my own moral commitments, for example, I still need to make a decision. This section builds on the work of Sen and Levi regarding rational choice under unresolved conflict of interests, and the further interpretation of their work by Gaus. Section 3 discusses in more detail in light of these considerations how it may be the interest of all reasonable citizens that there should be an effective problem-solving agency. This leads to a normative principle of political justification I call *pragmatic robust instrumentalism* (PRI).

Section 4 then moves on to epistemic theories of democracy more generally, and examines the general structure of that type of argument, and the assumptions and requirements of such a form of argument on the basis of a general theory of political legitimacy. In this section I concede that the *standard* versions of the epistemic argument for democracy indeed face an insurmountable problem given the facts of disagreement

and uncertainty. The standard version of this argument relies on what I call the *juror model* of epistemic reliability, which equates the epistemic task of political procedures with *truth-tracking*, that is, the likelihood of getting a given decision right, or equivalently, a relatively high average ratio of true to false decisions across all decisions taken. The problem with this is that the uncertainty assumption precludes the claim that any particular procedure (democratic or otherwise) can be expected to deliver this truth-tracking result.

Section 5 offers an alternative way of justifying political authority of democracy, namely as a *method of inquiry*. I conclude therefore that any realistic justification of democratic authority that takes the importance of outcomes into account must be a *methodological* one that understands political procedures as *progressive methods of inquiry*, and hence that we should focus (in evaluation as well as practice) on the aspects of democracy that are *functional for such a method of inquiry*. Finally, section 6 translates the pragmatic model of problem-solving into concretely *political* terms. This section fleshes out this pragmatic understanding of universal reliability in problem-solving: to the extent that they functionally enable experimental problem-solving in the policy-process, democratic decision-making mechanisms enjoy justified legitimacy. This argument, as will be shown, is *robust* with respect to the assumptions made regarding the definition of “good decisions,” and the form of the argument fulfills both the rationality and optimality requirements.



## 2. The Possibility of Rational Choice Under Uncertainty

At least since Rawls' so-called *political turn*, the *fact of reasonable disagreement* has become the fundamental fact delimiting legitimate state action and legitimate government. This is the basis of the entire project of *justificatory liberalism*, or the “public reason” project more generally: Which policies, laws, and forms of rule can be justified given the fact that reasonable people disagree about their value?<sup>110</sup>

For the purposes of this work, the most important aspect of the justificatory liberalism project refers to the justification of democracy. In particular, as several scholars point out, because reasonable disagreement *implies* that there is no external standard of the quality of decisions, any justifications of laws based on their quality are *prima facie* suspect. For this reason, laws and policies have to be justified not on the basis of their content, but always on the basis of the *procedure* by which they have been chosen, which in turn is justified through public reason, i.e. reasons that anyone could or would endorse, as long as they are reasonable.

This “flight from substance,” as David Estlund calls it, remains an important basis of the claim that there can be no substantive justifications of democracy.<sup>111</sup> The fact of disagreement is frequently taken as a fundamental objection to the idea of a substantive theory of political authority in general. Examples are the works of, among others, Charles

---

<sup>110</sup> see for instance Gaus, *Justificatory Liberalism : An Essay on Epistemology and Political Theory*; Gaus, *The Order of Public Reason : a Theory of Freedom and Morality in a Diverse and Bounded World*.

<sup>111</sup> Estlund, *Democratic Authority: a Philosophical Framework*.

Beitz, Jeremy Waldron, Fabienne Peter, Thomas Christiano, Nadia Urbinati, and Sean Ingham.<sup>112</sup>

This section has two aims, therefore: establish that we can speak of better and worse decisions under circumstances of deep disagreement, and establish that the appropriate response to disagreement is not the acceptance that there are no extra-procedural standards that indicate better or worse decisions, but trying to making an intelligent choice *despite* the disagreement that might eventually overcome that disagreement. If this latter response can be shown to be in “everyone’s interest,” then there may be a sense in which there is a “common good” in the political sense, and deep disagreement should prompt us to search for it, rather than conclude that it is inaccessible.

To put it differently, we should understand the disagreement with respect to political questions as an *epistemic* question, and in particular as a “weak” epistemic question in Robert Talisse’s sense.<sup>113</sup> This means that disagreement presents us with a problem of uncertainty that is in principle resolvable, even though the solution is not obviously available as things stand right now. This can be distinguished from a “strong” epistemic understanding of disagreement, which conceded that there may be a possible resolution to the problem but that it is in principle inaccessible to us humans.<sup>114</sup>

---

<sup>112</sup> Beitz, *Political Equality*; Waldron, *Law and Disagreement*; Fabienne Peter, *Democratic Legitimacy, Routledge studies in social and political thought* (New York: Routledge, 2009); Christiano, *The Constitution of Equality : Democratic Authority and Its Limits*; Urbinati and Saffon, “Procedural Democracy, the Bulwark of Equal Liberty;” Sean Ingham, “Disagreement and Epistemic Arguments for Democracy,” *Politics, Philosophy and Economics* (2013).

<sup>113</sup> Robert B Talisse, *Pluralism and Liberal Politics* (New York: Routledge, 2012).

<sup>114</sup> And this must yet be distinguished from the “ontological” thesis that there *is* no meaningful resolution to these problems of disagreement.

Initially, the way to assert the existence of a “common good” standard by which to evaluate policy decisions (which seems to be the precondition for a robust instrumental theory such as I propose) would be to posit a substantive ideal that goes beyond individual preferences. For instance, we might think that *justice* or *equality* are substantive ideals such that we can evaluate policy decisions with respect to how they advance them, regardless of whether anyone actually believes in (that specific form of) justice or equality. However, this of course runs afoul of the robustness requirement that is an essential part of *robust* instrumentalism. Epistemic justifications of democracy, and reliability claims, cannot be premised on a substantive understanding of the truth. However, of course this does not rule out that these ideals *may* actually hold. This position merely holds that they may not be used in the context of a justification.

So, this section takes an initial cut at this problem, without assuming a *preference-independent* standard of the common good. As Amartya Sen and a number of social choice theorists after him have pointed out, we may speak of *rational* collective decision-making even without assuming a definite standard of the quality of decision-making beyond the satisfaction of individual preferences. So the argument in this section is not to metaphysically rule out the existence of independent common-good-standards, but to suggest some ways to think about the existence of the common good *even without* assuming such standards. In the next section I will then go beyond brute preferences. The upshot in any case is to establish a possibility result, in order to support the view that *disagreement should be the start of inquiry, not the end.*<sup>115</sup>

---

<sup>115</sup> I will also return to this general topic in chapter 4 below.

### *Reasonable disagreement*

Let me start therefore from the *fact of reasonable disagreement* about politics: reasonable people actually do disagree quite deeply about what policies should be pursued: and they seem to disagree both in terms of the goals they want to achieve, and if they should agree about the goals, there frequently is disagreement about the causal relationships between policies and outcomes as well. As Elster puts it, there is a *double indeterminacy*.<sup>116</sup> Additionally, we have to assume that it is not obviously the case that one of the parties has made an inadvertent error of rationality (which could easily be corrected), or has morally unacceptable beliefs (e.g. psychopathic ones). Reasonable disagreement means that even after a period of reflection and a reasonable amount of information-seeking, people may disagree in their evaluations of policy. And this disagreement may not be rationally resolvable, given the present state of information.<sup>117</sup>

This is the baseline assumption that is shared by most current theorists of political legitimacy, and it seems to me indisputable that a degree of reasonable pluralism of opinion exists in most societies. What I want to contest in this section is the view that this disagreement entails that there are *no* standards by which to assess collective decisions. Let us call the conclusion drawn by the proceduralists the *no-standard-thesis* (NST). This states that in cases of ex ante irresolvable conflict over political questions, there is in principle no independent standard by which we may distinguish better from worse decisions.

---

<sup>116</sup> Elster, *Securities against Misrule* ch. 1.

<sup>117</sup> That is, there may not be enough reasons available (yet) that reasonable people could respond to, in order to resolve the disagreement. This lack of reasons is due to fundamental *uncertainty* about political questions—whether normative or factual/causal.

So in what sense could we possibly speak of better and worse policies or outcomes in the face of such unresolved disagreement? First, without going deep into this, we may ask how pervasive this disagreement really is. First, there seems to be little (reasonable) disagreement, at least within most liberal industrialized countries, regarding a quite extensive list of basic human rights. Rights to life, liberty and some degree of property, civil and political rights, rights to shelter and sustenance seem to be relatively uncontroversially held to be good things. We can probably safely assume that there is widespread agreement on a list of “primary bads” such as Estlund identifies: things like war, civil war, famine, epidemics, economic and political collapse, are pretty much universally identified as bad things.<sup>118</sup> Beyond this, there seems to be quite widespread agreement about the *pro tanto* value of certain political outcomes: economic growth, environmental integrity, public health, external security.<sup>119</sup> And even with respect to modes of life, while there is clearly no agreement as to what a good life consists of, there seems to be some agreement on what does *not* constitute a good life: a life of substance abuse, for instance, or one spent without the opportunity to exercise one’s capacities at all (due to unmitigated disability, poverty or “unemployability”).<sup>120</sup> Furthermore, disagreement about ideal ways of life does not necessarily lead to equal clashes over policy. We may disagree utterly with respect to which way of life we hold valuable, while completely agreeing that there should be no policy enforcing any particular way of life. If

---

<sup>118</sup> With some possible exceptions, of course. See my discussion of anarcho-libertarians and their positive view of political collapse in the previous chapter. For the idea of primary bads generally, see Estlund, *Democratic Authority: a Philosophical Framework*, ch. 9

<sup>119</sup> This means that, *ceteris paribus*, more of those things is better than less.

<sup>120</sup> Sure enough, there are some people who value those ways of life, and there may be some people who are against environmental integrity, but the question is whether those really are instances of reasonable disagreement.

disagreement has this structure, of course the *good* political decision is the one that guarantees neutrality between ways of life. So here are already some considerations about when disagreement does not seem to entail the lack of political standards (even though there may be a lack of comprehensive *moral* standards—as to the question of the good life, for example).

There are, of course, genuine differences with respect to values as well: some people value equality of resources, while others do not value this at all. Some people value other people's intimate relationships only in their heterosexual variants, others rate them all equally as long as they are loving ones. Furthermore, there is of course quite extensive disagreement as to the *relative ranking* of all the values above: we may have to trade off economic growth with environmental integrity, or the protection of individual property with economic efficiency. Nevertheless, if we conceive wide parts of reasonable disagreement as one of conflicting *rankings* of what are agreed to be *goods*, and not blunt conflict of different values, we may already see some potential for identifying what a “good solution” may be: *partial rankings* may be possible if we do not have to resolve brute clashes of preference. Finally, there is of course quite deep disagreement about what kind of intervention will *realize* those goods: how to achieve economic growth, for instance. Do we achieve that through increased spending or spending cutbacks? The disagreement not only between “ordinary citizens,” but also between economists on this question illustrates the depth of the divide.

However, this example also indicates that disagreement does not always imply that the NST holds: we should not abandon the goal of economic growth *purely* because there is disagreement about how to get there. The important question is of course

whether we can see what appears to be *value conflicts* in the same way as admitting of better and worse solutions. I will argue that we can think of this in the same way.

So, the extent of deep reasonable disagreement may not encompass the entire political field, but there are still many deep disagreements. The question is whether this residual disagreement is so large that we are not warranted to speak of “objectively” better and worse political decisions. The key issue is of course how we understand “reasonable” disagreement. This can be defined more or less “thickly” as Rawls’ put it: and the possibility of rational choice under disagreement depends crucially on the sense in which the disagreement is “reasonable.” In the following I will suggest how we may understand “reasonable,” first in a very thin way, and later in a somewhat more substantive way.

#### *Rational decisions and unresolved disagreement*

So, in spite of deep disagreement, and crucially, also without necessarily resolving the (reasonable) disagreement, the NST does not follow directly from the fact of pluralism. There is a *possibility* of better and worse decisions without adopting one final perfectionist standard. As I have already put it in the last chapter, a good political decision should be seen as one that answers the question *what is to be done*, not (necessarily) the question of ultimate normative value.

For the purposes of this initial argument I shall adopt Sen’s conception of a “maximal set.” The maximal set is the set of those decisions that are jointly optimal: there is no rational way to decide which of the options in the set is better than any other one

within it, but we can clearly decide that any option within the set is *better* than any one outside of it. The question is: how big is the maximal set under conditions of reasonable disagreement? If it is relatively small, and especially if it does not include the status quo, then while we may not be able to uniquely identify an optimum, we have a useful standard of quality for wide areas of the policy spectrum.<sup>121</sup>

There are a number of cogent arguments by which we can reduce the size of the maximal set even under deep disagreement. This indicates that while *some* political problems may remain rationally irresolvable (think of the ultimate value commitments underlying the abortion debate), there are great areas for possible progress in policy: opportunities for moving into the maximal set from somewhere outside. In this vein, a procedure might be justifiable which does the right thing *despite* the disagreement.

First, let us consider a clash in rankings. One may think that despite disagreement, *non-dominated* outcomes are better than dominated ones. In other words, a good decision procedure should at least not bring about outcomes that are clearly dispreferred by everyone in comparison to its alternatives. Consider the following case:<sup>122</sup>

*Case 1:*

Party 1:  $A < B < C$

Party 2:  $A < C < B$

Here, the two parties disagree in their relative ranking of B and C, but whether or not C or B is the eventual outcome, both parties prefer this to A. So while in this situation we

---

<sup>121</sup> Note that a small maximal set is “better” than a large one, for the purposes here.

<sup>122</sup> “<” is taken to mean “is strictly dispreferred to,” and the reverse for “>”



cannot say which choice is the uniquely best one, it is clear that the set {B; C} is *better* than A. Note that there is *disagreement* about ultimate ends, but still an (incomplete) standard available.

Before this is dismissed as a merely theoretical possibility, let me point out that the *dominated* set may be quite large; and that many existing political systems frequently do end up stuck with dominated outcomes. One might think, for instance, that at the very least policies that lead to civil war or generally to the breakdown of state functionality are dominated in this sense. So, even if there is reasonable disagreement, there is no-one who prefers civil war to any non-civil-war option. Following Sen, then, we can call the set {B; C} of non-dominated options the *maximal set*. So, despite all the disagreement, a procedure that chooses either B or C over A is clearly *better*.

An example of such a structure may be the recurring budget debate around the “debt ceiling” in the United States Senate. The options, simplifying somewhat crudely, are government default (A), a high-spending, high-tax budget (B) and a low-spending, low-tax budget (C)—and the Republicans and Democrats may be represented by party 1 and 2 respectively. Clearly, a political system that tends to avoid default is pretty good, even though we might not be able to say whether C or B would be “better” decisions.

Now, the one thing that anyone who has followed the debates in 2012 and 2013 can notice is that Party 1 adopted a bargaining position and claimed that their actual preferences were these:<sup>123</sup>

Party 1':  $B < A < C$

---

<sup>123</sup> this is another instance of the benefit of appearing to be reckless or "crazy" in some strategic situations. See Thomas C Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1980).

Whether or not this was actually a misrepresentation; we might agree that it is a virtue of a political system (perhaps even a necessary legitimacy condition) that in cases like *Case 1*, it should not produce or encourage misrepresentation such that we end up with an dominated outcome that nobody wanted. This possibility of strategic misrepresentation illustrates another point: that some *apparently* ultimate preferences may be unreasonable because they are not genuine; and we cannot draw the NST conclusion unless we can expect to rule out such misrepresentation.

Nevertheless, assuming party 1 was sincere, this would make it impossible to restrict the maximal set to B and C. This may happen when two parties are completely opposed in their outlooks. So now, let us consider a case with exactly opposed ideal points. Here, however, disagreement at the transcendental level may imply partial agreement on the practical level, even if parties to the disagreement may not see it that way (and may well *de facto* also claim disagreement about that). Consider the following example of two parties and their preferences.

*Case 2:*

Party 1:  $C < SQ < B < A$

Party 2:  $A < SQ < B < C$

Let's assume these describe the rankings of relative states of affairs of two fundamentally opposed parties. Whatever is ranked highest by party 1, is ranked lowest by party 2. However, there is obviously partial agreement that B is to be preferred to the status quo SQ. Therefore, it seems to me, the ranking implies that there is clearly a truth about what should be done *in the face of radical disagreement*: we should try and bring about B. If party 2,

then, would claim to prefer SQ to B, they must be either mistaken about their preferences, have perverse preferences (for example, “even though I would personally like to live in status B, it is even more important for me to harm the other guys”), or deliberately misrepresenting them in order to gain a bargaining advantage. The former two possibilities would mean that the disagreement at the practical level is of no moral relevance, while that latter possibility will occur *only in a system which rewards such misrepresentation*.<sup>124</sup> Again, therefore, in that case it is questionable whether the dishonest disagreement is also morally relevant. In my view, such misrepresentation would count against the use of such systems, regardless of their other procedural virtues.

If, as described above, we view the problem of disagreement as an *epistemological* problem of *uncertainty*, we can easily evaluate procedures according to the quality of the *decisions they make under (or despite)* that uncertainty. This need not be the same as the capacity to correctly resolve the disagreement; under uncertainty epistemic reliability could mean the capacity to make all the improvements one can. In the terms of the example, a procedure that would tend to move to B could be justified on that basis, even though it would not resolve whether A or C is better. A procedure that would reveal the truth preferences of the parties would, in this case, be strictly (as well as objectively) preferable to one that induces them to strategically misrepresent them.

In addition, an epistemic justification can, and ought to, include considerations of *de facto* disagreement as well. It is a fallacy to assume that instrumental justifications generally cannot include procedural considerations. Those arguments are only committed

---

<sup>124</sup> That democratic procedures can be subject to this is a key normative implication of the Gibbard-Satterthwaite Theorem (see chapter 6 below).

to evaluating procedural aspects in terms of the instrumental value they have.<sup>125</sup> For instance, if a procedure should choose an “right” policy under conditions of extreme disagreement, this might lead to widespread discontent—to noncompliance and perhaps unrest; eventually therefore, such a procedure would not lead to the best outcomes overall. In those cases, the substance of the decision might have been *pro tanto* “correct,” but choosing the policy was not the right thing to do, all things considered. In so far as citizens’ reaction or compliance is part of the objective context in which a decision is implemented, it should influence the rightness of that decision. Instrumentalism does not commit one to a “fiat iustitia et pereat mundus” attitude.

Somewhat related to this idea is that whatever decision mechanism we use to reach a decision under unresolved disagreement, at least it should not involve corruption, misrepresentation, systematic bias, or other forms of self-serving abuse of the system. This may be a *procedural* limitation on what can be considered a good decision under disagreement. This reasoning forms the basis of Jon Elster’s recent idea of a *negative* procedural theory of collective decision-making.<sup>126</sup> We can try our best to eliminate these pathologies of collective decision-making, even though we might not be able to say anything *positive* about which solutions are in the end better than any other.

---

<sup>125</sup> Daniel Viehoff, “Debate: Procedure and Outcome in the Justification of Authority,” *Journal of Political Philosophy* 19.2 (2011): 248–259; Dworkin, “What is Equality? Part 4: Political Equality;” Kasper Lippert-Rasmussen, “Estlund on Epistocracy: A Critique,” *Res Publica* 18.3 (2012): 241–258.

<sup>126</sup> Elster, *Securities against Misrule*.

*Another criterion*

We can exclude even more outcomes from the maximal set by introducing the idea of *cardinality* in the different possible rankings. For instance, we might rank the options

$$C < B < A,$$

but might think that B is only minimally better than C, while A is a lot better than B. As it turns out, if we allow for this possibility, the maximal set may become even smaller, and the rational choice even more determinate.

Here, Isaac Levi's notion of *V-admissibility* provides valuable insight. This is a criterion to denote a set of admissible rational decisions under unresolved conflict over different sets of valuations.<sup>127</sup> The basic idea is as follows: In a situation of choice between a finite number of options, we can represent different moral commitments as functions assigning different cardinal values to each option. In other words, every commitment we hold cardinally ranks all available options. Now, we may hold several moral commitments ourselves, but we still have to make a decision. But we can also understand this logic in the context of interpersonal disagreement, where a collective decision must be reached despite unresolved differences in valuation between persons. On the basis of this model, then, *political* disagreement is based on the fact that people have different value functions that assign different values across all the possible options.

Now, in our decision process we can differently weight the different value-assignment functions (what I have called commitments). Any decision to reach a collective

---

<sup>127</sup> Isaac Levi, *Hard Choices : Decision Making Under Unresolved Conflict* (Cambridge: Cambridge University Press, 1986).

ranking under these conditions (including the decision to remain in the status quo) is such a weighting. A dictator for instance, assigns zero weight to all other value-assignment functions, and a weight of 1 to his own. Majority rule, on the other hand, gives a different weight to the different rankings. Let us assume that in collective decision-making a permissible weighting of different positions should at least assign a non-zero weight to every person's ranking. Every set of weights represents one potential *resolution of the conflict*. In other words, for each option we can take the weighted average of everyone's value assignment, yielding a cardinal value for each option and for each possible set of weights.

The NST assumption in this context implies that that no uniquely acceptable set of weights (no uniquely acceptable resolution of the conflict) exists, and therefore all options are *prima facie* equally valid—we should just choose a procedure *ex ante* and then settle on the outcome as it turns out.

Now, the key outcome of seeing decision-making under conflict in this way is that *even though* we cannot decide between the different resolutions of the conflict, it may turn out that some options are not optimal under *any* potential resolution (set of weights). Therefore, even though we may not be able to resolve the conflict, it would be irrational to choose any option that which no acceptable conflict resolution would pick.

This is a complicated way of putting some very intuitive notions. Consider for example the following options and valuations in Case 3 below.<sup>128</sup> In this situation, Option B is not actually dominated—it is part of the maximal set together with A and C. Indeed, there is no Pareto-efficient way to move to B if we are at either A or C. So it seems there

---

<sup>128</sup> This is adapted from Levi's own example, *ibid.*, 11-3

is no principle to choose between A, B or C. However, common sense suggests that B should not be chosen. It is only marginally better than the worst payoff in either other case, while the added benefit of choosing either A or C is pretty big.

*Table 1: Case 3*

|         | Option A | Option B | Option C |
|---------|----------|----------|----------|
| Party 1 | 100      | 51       | 50       |
| Party 2 | 50       | 51       | 100      |

Now, if we look at this closely, we can see that there is no permissible set of weighted averages of the benefits accruing to 1 and 2 such that B would come out on top. However we relatively value the respective benefits, B is never optimal. Weighing the interests of 1 and 2 equally at .5, for instance, either A or C would be collectively valued at 150, and B at 102. The more we weight 1's interest, the more attractive A becomes relative to B, and the more we weight 2's interest, the more attractive C becomes. Unless we absolutely require Pareto efficiency, B will under no valuation turn out optimal. In Levi's terms, B is not *v-admissible*.

This reflects a very basic intuition. We can also come up with a fitting political example. Consider the following situation of conflict over economic growth. Our example country contains only two types of people: Keynesians and free-market enthusiasts, and the three possible options are: High-Tax High-Spending, Low-Tax High Spending (this is the Status Quo), and Low-Tax Low-Spending. Now assume the preferences are as follows:

Table 2: Case 3a

|                | High-Tax,<br>High-Spending | Low-Tax, High<br>Spending | Low-Tax, Low-<br>Spending |
|----------------|----------------------------|---------------------------|---------------------------|
| Keynesians     | 100                        | 51                        | 50                        |
| Free-Marketers | 50                         | 51                        | 100                       |

Keynesians of course like the expansionary state, and do not like leaving investment decisions only to the free market, and for the free-marketers the situation is the reverse. However, for the Keynesians, the status quo is only little better than the free-market situation, since it is extremely inefficient and piles up debt. The free-marketers equally judge the status quo only little better than the Keynesian state: there may be lower taxes, but as persons with rational expectations they fully expect that in the end they will have to pay up.

Now, in this situation, remaining in the status quo makes nobody happy. According to *v-admissibility*, however, either one of the clear-cut options is preferable to the middle way. We might think that under unresolved conflict, we may therefore restrict the “good choices” to the intersection of the maximal set and the *v*-admissible set. A political procedure that in this situation would settle on Low Tax High Spending would be sub-optimal.

As we can see, there are a number of ways to understand better or worse solutions under situations of unresolved conflict: there are a number of ways to think that even



though there may be some residual disagreement about ultimate ends, we can find some criteria that rejects the NST.

Of course, interpersonal comparisons of utility (especially in a cardinal sense) are deeply problematic. Indeed, we may never know whether we are in a situation like Levi envisaged; that may be epistemically impossible. However, my goal here was only to suggest that it is possible that political conflict may be of this kind, and more importantly, given this possibility, the NST advocate must adopt a strict standard of Pareto-efficiency to rule out the *v-admissibility* argument. This illustrates that the NST is not as immediately obvious as it may initially seem. In order to rule out *v-admissibility*, the defender of NST must adopt strict Paretianism as a normative premise.

### 3. The Public Interest and the Common Good

Finally, I shall turn to another argument narrowing the maximal set. A version of this is found in Gerald Gaus' recent writings.<sup>129</sup> Philip Pettit and Christian List have also made an argument in a similar vein regarding how we can think of a common standard for political decision-making under conditions of disagreement.<sup>130</sup> The basic argument is this: If we define "reasonable" in a slightly more substantive sense—more substantive than strategically seeking one's own self-interest in every situation—then we can come up with

---

<sup>129</sup> Gaus, *The Order of Public Reason : a Theory of Freedom and Morality in a Diverse and Bounded World*.

<sup>130</sup> Philip Pettit, "The Common Good," in *Justice and Democracy: Essays for Brian Barry* (Cambridge: Cambridge University Press, 2004); List and Pettit, *Group Agency : the Possibility, Design, and Status of Corporate Agents*; Philip Pettit, *On the People's Terms* (Cambridge: Cambridge University Press, 2012).

a maximal set that is more determinate. This is what Gaus calls the “socially eligible set.”<sup>131</sup>

*The alternative to agreement*

The key is that reasonable agents should recognize the important social coordination role played by a unique set of social norms that is generally agreed upon and which is endorsed and (coercively) enforced within a community. There are substantive efficiency gains from agreeing on a set of such norms that arrange social behavior—if only it avoids constant conflicts over day-to-day-issues.<sup>132</sup> Justified moral norms are therefore those that can be endorsed by reasonable people (“members of the public” in Gaus’ terms) who, among other things, recognize the coordinative value of having a shared and agreed-upon set of norms—and perhaps a state based on them—in the first place.

Therefore, paraphrasing somewhat, when members of the public comparatively rank different policy proposals, reasonable individuals keep in mind the cost of *failing to agree* on any policy at all. They are not Schelling’s bargainers who want to appear unreasonable and reckless to gain a bargaining advantage. The relevant comparison is not only between the policy proposals that happen to be on the table, but between the proposals and the failure to coordinate behavior at all.<sup>133</sup> We can see that this is another

---

<sup>131</sup> Note that “reasonable” is still not defined “thickly,” in Rawlsian terms, that is, according to a comprehensive doctrine.

<sup>132</sup> This Humean understanding of the role of norms is also emphasized in Philip Kitcher, *The Ethical Project* (Cambridge, Mass.: Harvard University Press, 2011).

<sup>133</sup> Gaus, *The Order of Public Reason : a Theory of Freedom and Morality in a Diverse and Bounded World* ch. 17.

avenue that potentially narrows down the maximal set. Especially, there is pressure that pushes *toward a single coordinating solution* to policy problems, even though we might not be able to conclusively say which solution that should be.

We may also understand this idea in terms of “interest.” As Barry points out, a policy can be considered to be in someone’s interest only in comparison to an alternative.<sup>134</sup> This also means that a reasonable person seeking to advance their interest would not assert their brute preferences, such as their absolute ideal. They assert their preference according to what is in their interest given the alternatives. Sometimes apparently irresolvable disagreement may be due to a failure to appreciate the actual alternatives. As Barry puts it, disagreement may be due to people holding “secret alternatives;” think of people constantly comparing the actual options on the political table with a ex-post rationalized and idealized version of the “Good Old Days.” Of course, nothing will ever do in comparison with that.

The fact of brute disagreement *in fact* does not, therefore, necessarily indicate a conflict of interests. In Barry’s words, “To point out as if it were a great discovery that all proposals *which are actually put forward* meet opposition is as naïve as expressing surprise at the fact that in all cases which reach the Supreme Court there is something to be said on each side. (If there isn’t, someone has been wasting an awful lot of money).”<sup>135</sup>

Thus we can say that when thinking about a standard of decision-making under disagreement we may focus only on what Pettit calls “avowable” interests, rather than

---

<sup>134</sup> Brian Barry, *Political Argument* (London: Routledge & Kegan Paul, 1965) 194.

<sup>135</sup> *ibid.*, 195.

brute representations of interest.<sup>136</sup> Thus, for example, no *particular* road to be built is of course in everyone's interest. Why should I support a road being built in a part of the country I will never visit? But it may be in everyone's interest that there is a system in place that builds roads when it should be necessary. And this should be a system that includes a procedure to reliably determine when a road needs to be built, and, very importantly, efficiently gets said road built.<sup>137</sup> Thus, to the extent that we can be held to have this second-order interest, a first-order interest with respect to a particular road (or policy) does not necessarily imply the same level of disagreement at the second-order level.

More generally, it may be assumed that as members of a political community, we have a common interest in rules that coordinate collective behavior<sup>138</sup>, that also include rules for choosing political actions that are consistent with a common interest we have in such a functioning system.<sup>139</sup>

*Pragmatic robust instrumentalism*

In other words, we should understand the problem of political collective action not as one where different people's "private net interests" conflict, but one where everyone's net interests are addressed *within the framework of an organized system*. As Pettit puts it, "It is in the

---

<sup>136</sup> Pettit, "The Common Good."

<sup>137</sup> Barry, *Political Argument* 197.

<sup>138</sup> Gaus, *The Order of Public Reason : a Theory of Freedom and Morality in a Diverse and Bounded World*.

<sup>139</sup> Barry, *Political Argument* 201; Pettit, "The Common Good" 156.

avowable net interest of every citizen that there is a state that is forced to track the common good.”<sup>140</sup>

Since the content of the concept “common good” is of course *uncertain* (as has been discussed at length), I propose we should see this as follows:

*Pragmatic Robust Instrumentalism (PRI):*

It is in everyone’s interest that there should be a political system that can be expected to and is robustly capable of, solving common political problems, as they arise under conditions of uncertainty, and whatever the solution may be.

In other words, it is in everyone’s interest that there should be a political mechanism in place that *reliably* resolves political problems, *whatever will work* as a solution. If we can make a case that democracy is reliable in this sense, it can be justified. Note that this is something even the fringe libertarians from chapter 2 can endorse: in so far as they have an interest in living in a community with others, where there are bound to be conflicts over policy (or the need for policy), they have an interest that those conflicts (or “problems”) are resolved. Even though they themselves endorse a particular way to resolve these problems (the libertarian way), given the fundamental uncertainty about this they have *practical* interest in that the procedure should reliably resolve these issues.<sup>141</sup>

Now, putting all these considerations together, we can see that even under the assumptions of justificatory liberalism and the fact of disagreement, we can possibly speak of an indeterminate set of “common good,” or objectively “better” solutions. The

---

<sup>140</sup> *ibid.*, 156.

<sup>141</sup> Recall the discussion in chapter 2, section 6.

principle of PRI focuses on the “methodology” of the political decision mechanism, not the final outcomes it achieves. In the words of the previous section, we focus on moving into the maximal set from without, not on which of the elements of the maximal set we should choose. Therefore, it focuses on problems and failures from which we move *away*, not a goal *towards* which we move.<sup>142</sup>

That is, whatever my “brute” private preferences are, as a reasonable person I should realize and endorse this PRI principle. Therefore, a political system may be justified on the basis of the principle PRI *despite* the fact of disagreement and uncertainty. Of course, the simple models discussed in this section are problematic since we typically do not *know* what the profile of preferences and valuations actually is. However, what this section tried to show is the *implausibility* of the NST in its strong epistemic form—at least as a direct inference from the fact of disagreement. Deriving NST from disagreement requires another normative premise which rules out any of the considerations just presented (for instance, strict Paretianism).

This uncertainty also highlights my general point: the reasonable disagreement we observe should prompt us to *further investigation* into the problem. Disagreement ought to be seen primarily as an *epistemic* problem. Think back to the last example: we don’t know whether the profile of Case 3a obtains; but it might, and if it does, it would be beneficial to find out. To restate the point made above: *disagreement should be the starting point for further inquiry, not the end point.*

---

<sup>142</sup> See for example Wiens, “Prescribing Institutions Without Ideal Theory.”

This section has only established in abstract terms how a standard of the quality of democratic (and more generally, political) decision-making can be understood. The abstract principle of *Pragmatic Robust Instrumentalism* has been introduced as a basis for an argument for outcome-based democratic legitimacy under conditions of deep disagreement and uncertainty. The next sections move toward the question of how this principle and its consequences can be understood more concretely.

#### 4. Universal Reliability and How to Argue For It

So, even if we now have a modest idea of a standard for the quality of political decisions even in the face of apparently irresolvable disagreement, there is the question how we can support a claim that democratic procedures will result in decisions of that quality, regardless of what that quality is. How, as it were, should we *operationalize* the philosophical principle in terms of democratic decision-making? As mentioned above, we have to make a claim for the reliability of democratic procedures that fulfills the PRI principle.

The simplest for of this claim would be for us to show that we have reason to believe that democracy will produce more good decisions overall than its alternatives, or that it is more likely to get things right than others. This is probably the “standard” operationalization of epistemic arguments for democracy, namely that the capacity or tendency of democracy to produce good outcomes, should be understood in terms of its *truth-tracking capacity*, or the likelihood, for a given decision, that a democratic procedure will get it right. If a common good exists (as postulated above), these justifications argue,

democracy will tend to find it more reliably than its alternatives. However, as I will show, there is a basic problem with using such an argument in the context of a *robust* justification of democracy; we will have to ground democracy's instrumental value in some feature other than truth-tracking reliability. I will present my own alternative conception in Section 5 below.

Recently, epistemic arguments for democracy have enjoyed a certain renaissance. They rely on a claim that democracy has an epistemic “edge” over other political decision mechanisms, an edge when it comes to making the *objectively right decisions*. However, as we already mentioned we may not *substantially* define the right decision; that would violate the robustness requirement. Epistemic justifications of democracy are grounded in a claim that democratic procedures can reasonably be expected to have an epistemic edge over other procedures *whatever the truth might turn out to be*. For that reason, these justifications have to be based on asserting the *universal* reliability of democratic procedures. However, asserting the *universal* reliability of democracy also seems problematic: that regardless of what the truth is, democracy is likely to find it and decide accordingly.

I argue in this part of the chapter that usual versions of the universal reliability argument essentially are too demanding. They cannot resolve the dilemma outlined in chapter 2—that we care about outcomes, but do not know how to assess them *ex ante*. The consequence is a wrong focus of the debate about “epistemic reliability”: advocates as well as critics tend to define this as maximizing the *likelihood*, for a given decision, to decide correctly, or alternatively as the *ratio* of right to false decisions across the board. A reliable procedure, in this sense, “tracks the truth.” This models the epistemic task of



political institutions essentially on that of the *judge* or *juror*, delivering verdicts with different degrees of reliability. However, it is impossible to argue that democracy—or any other procedure—would maximize this ratio without at the same time defining what “correctness” means in the first place. To put it differently: we cannot make convincing arguments for the truth-tracking capacity of democracy that are *robust* across different scenarios of what the truth will turn out to be.

However, all is not lost. We are not tied to the juror model of understanding reliability. In fact, outside the courtroom, we do not generally think of reliability in those terms. In other fields also characterized by great uncertainty about what the right answers might be—think of scientific inquiry, engineering, medical research—we do not understand “reliability” in this way. Scientists or engineers are not distinguished by a high likelihood that any one of their propositions will turn out correctly; we cannot know that beforehand. What is important is that the scientist employs the right method: the ability to generate possible solutions, and the capacity to identify what works and what doesn’t and to *adapt* accordingly. In this way she can *make pragmatic epistemic progress over time*.

Likewise, in the remainder of the chapter I argue that a universal reliability argument should be based on democracy’s *pragmatic* ability to experiment and gradually adapt to successes and mistakes. This is superior to arguing from its truth-tracking ability. Thus, we do not need to assert, controversially, that for a given decision, democracy is more likely to get it right than other decision. It is enough to argue that *as a procedure, democracy is systemically adaptive*. In contrast to other systems, it has a built-in mechanism for self-correction and adaptation.

For such a theory of democracy, it becomes relatively less important *who* is making policy decisions and indeed *how* they do so, as long as there is a properly working *feedback* mechanism that enables a form of social learning. As I will suggest, it is plausible democratic systems can provide this essential function, and we should see this pragmatic, evolutionary function as their key advantage. In contrast to the juror model the pragmatic version can pass the public justification test of rationality, optimality and robustness. Therefore, this pragmatic conception of the principle of robust instrumentalism can serve as the basis for a justification of democratic authority.

*The “juror model” and truth-tracking as the standard*

Let us look once again at the structure of what we may call the standard epistemic justification of democracy. The focus is on epistemic proceduralism, the idea that a procedure can be justified if we can reasonably expect it to be best among possible procedures at getting things right, even though in some particular instances it might not do so. On this basis, an epistemic-procedural justification of democracy has the following general structure:<sup>143</sup>

- (1) *Principle of Justification:* Those institutions are politically justified that are elements of the set of possible procedures, and are on the whole better than all other elements of the set at getting decisions right, across the set of decisions that they can be expected to encounter.

---

<sup>143</sup> This follows the form of argument in Estlund, *Democratic Authority: a Philosophical Framework*.

- (2) *Epistemic Reliability Claim*: Democracies can be expected on the whole to be better than all other possible procedures at getting decisions right, across the set of decisions that they can be expected to encounter.
- (3) *Conclusion*: Democratic rule is justified.

This leads to an obvious question: by which standard should we judge whether and when a procedure is “getting decisions right?” There are two ways to look at this idea: with respect to a given individual decision (“one-shot” synchronic reliability), or across a number of decisions (“long-term” diachronic reliability). We have to establish long-term reliability to ground claim (2). The standard interpretation of such a claim—espoused by advocates as well as critics—seems to be that this should refer to the ability of democracy to track the truth in its decisions, which conflates these two notions.

“For epistemic democrats, the aim of democracy is to ‘track the truth.’ For them, democracy is more desirable than alternative forms of decision-making because, and insofar as, it does that. One democratic decision rule is more desirable than another according to that same standard, so far as epistemic democrats are concerned.”<sup>144</sup>

Long-term reliability here is simply conceptualized as the aggregate of all discrete one-shot decisions, that is, as the ratio of right to false decisions over the universe of decisions made. Over a large number of decisions, this ratio is equivalent to the average likelihood to decide correctly for a given one-shot decision situation. In so far as this ratio is high, a reliable procedure therefore is said to “track the truth” through its decisions.

As Estlund points out, such a view of reliability comprises two aspects: sensitivity, that is the likelihood to judge  $P$  true if  $P$  is in fact true; and discrimination, the likelihood

---

<sup>144</sup> List and Goodin, “Epistemic Democracy: Generalizing the Condorcet Jury Theorem” 277.

that P is in fact true when P is judged true.<sup>145</sup> Equivalently, we can say a procedure's reliability is the likelihood that it judges a proposition P true if P is true and judges P false if it is false. In terms of relative frequencies this means that a procedure is a more reliable, the higher its ratio of correct over wrong judgments is—both on positive (“P is true”) and negative (“P is false”) judgments.<sup>146</sup> Given that political decisions involve normative as well as causal questions (“ends” as well as “means”), a procedure can be said to track the truth whenever it reaches high degrees of discrimination and sensitivity with respect to both normative and factual truths. Already we can see how demanding a standard this is: the aggregate of the people, organized democratically, must have a higher reliability on determining what the right or just outcome would look like, and on getting the relevant causal facts right, than a subset of the population.<sup>147</sup>

This conception resembles our idea of the reliability of judges or jurors. The perfect judge convicts all and only the guilty, and acquits all and only the innocent; the reliability of judges and juries ought to be measured by their approximation of that standard of sensitivity and discrimination. Hence I shall call this general understanding of reliability the juror model. The two key aspects of the juror model are, therefore, that long-term reliability is defined as the ratio of right over false decisions, and this ratio is equivalent to the average expected one-shot reliability across all decisions that will be faced in the future.

---

<sup>145</sup> Estlund, *Democratic Authority: a Philosophical Framework* 112-116.

<sup>146</sup> this definition is due to Robert Nozick, *Philosophical Explanations* (Cambridge, MA: Harvard University Press, 1981) 178.

<sup>147</sup> Of course, for moral or practical reasons we might put different weights on these two dimensions for certain contexts; especially if we cannot have a perfect procedure. Failures of discrimination, such as condemning an innocent man, might be given more weight than letting a guilty man go free.

Note that truth-tracking in this sense is just a formal standard of reliability that is prima facie independent of the status of the underlying notion of “truth.” The question is one of *normative* epistemology, that is, the question of what we think a valuable epistemic procedure (in this case a political procedure) should do. The answer to this might well be different in scientific, judicial, or political contexts. It is not a given that our epistemic goal should always be to indiscriminately maximize the truth-tracking ratio.<sup>148</sup> In any case, this question of the purpose of our epistemic procedure is conceptually independent of the epistemological question of whether the “truth” underlying epistemic performance is based on a realist, naturalist, contextual, relativistic or any other theory of truth. Recognizing this point, Estlund uses what he calls a minimal or deflationary account of truth, according to which a belief “x is F” is true if and only if x is indeed F, however that may eventually be filled in.<sup>149</sup> From this point on, my use of the word “truth” should be taken to refer to such a deflationary account of truth. This is in accordance with the model of robust instrumentalism, which demands that no substantive account of “goodness” should be postulated.

This conceptualization of epistemic reliability as a quantifiable likelihood or ratio facilitates comparative claims; that might also be the reason advocates as well as critics tend to adopt it, regardless of how they eventually fill in the meaning of “truth.” More generally, epistemic arguments implicitly assume something like the juror model

---

<sup>148</sup> For one thing, we could get this ratio up by asserting an endless number of true but completely irrelevant propositions. See Philip Kitcher, *Preludes to Pragmatism*, 32. In a political sense we may try to increase the number of pointless, but “correct” regulations.

<sup>149</sup> Estlund, *Democratic Authority: a Philosophical Framework* 25.

whenever they make the claim that decisions made by one group or procedure are more likely to be true than decisions by another group or procedure.

This truth-tracking ratio also tends to be the main bone of contention in debates about the benefits or drawbacks of large-group aggregative decision-making, for examples debates about the applicability of the Condorcet Jury Theorem. Among others, List and Pettit, for instance, also explicitly endorse this sense of reliability as the foremost epistemic desideratum of group decision-making.<sup>150</sup> In recent reconstructions of the CJT this remains the standard for assessing epistemic performance as well.<sup>151</sup> But also beyond the confines of interpreting the CJT, *truth-tracking* is taken as the appropriate standard by which to measure decision-making reliability, of democracy or otherwise. In defining the term “collective wisdom,” for example, Adrian Vermeule also takes truth-tracking to be the “baseline” desideratum of epistemic reliability, even though we might value some other things as well.<sup>152</sup> Hélène Landemore sums up this view of reliability as follows: “The sustained epistemic case for democracy that I propose in relation to this idea of democratic reason boils down to the simple following claim: democracy is a good collective decision-making procedure because, among other things and all things equal otherwise, it maximizes our collective chances to make the right choices.”<sup>153</sup>

---

<sup>150</sup> List and Pettit, *Group Agency : the Possibility, Design, and Status of Corporate Agents* ch. 4.

<sup>151</sup> Goodin and Spiekermann, “Epistemic Aspects of Representative Government;” Dietrich and Spiekermann, “Epistemic Democracy with Defensible Premises.”

<sup>152</sup> Adrian Vermeule, “Collective Wisdom and Institutional Design,” in *Collective Wisdom and Institutional Design*, ed. Jon Elster and Hélène Landemore (Cambridge: Cambridge University Press, 2012), 338–368.

<sup>153</sup> Hélène Landemore, “Why the Many Are Smarter than the Few and Why It Matters,” *Journal of Public Deliberation* 8.1 (2012): 3.

The reverse is also true: Critics of epistemic democracy tend to focus on claims for democracy's reliability when it is understood in this way, and their implausibility. Thomas Christiano and Gerald Gaus, for instance, focus on the claim (made by Estlund in particular) that democracy is more likely to make right decisions than a "random" procedure, and on the difficulty of interpreting and defending such a claim.<sup>154</sup> Fabienne Peter's and Sean Ingham's respective objections to epistemic justifications, on the other hand, are based on the argument that this comparative likelihood claim conflicts with plausible intuitions about the normative relevance of disagreement in democratic societies.<sup>155</sup>

In what follows I will side with the critics in denying that we can make a plausible claim that democracy is good at tracking the "truth" in its decisions. As I argue below, the "robustness" part of robust instrumentalism rules out arguments of this sort. *However*, in support of epistemic democrats I argue that we *can* make a plausible claim for democracy's universal reliability *without* having to make a claim about this truth-tracking likelihood. In order to do that, however, we have to interpret the task of political procedures *pragmatically* rather than on the juror model.

### *The Uncertainty Objection(s)*

The fundamental problem with the "juror model" is the following: in order to ascertain a procedure's expected long-term reliability ratio, we have to form an expectation of the

---

<sup>154</sup> Christiano, "Debate: Estlund on Democratic Authority;" Gaus, "On Seeking the Truth (Whatever That Is) Through Democracy: Estlund's Case for the Qualified Epistemic Claim."

<sup>155</sup> Fabienne Peter, "Pure Epistemic Proceduralism," *Episteme* 5.01 (2008): 33–55; Ingham, "Disagreement and Epistemic Arguments for Democracy."

procedure's average one-shot reliability across the entire set of at least all reasonably likely decision problems. This is a consequence of grounding long-term reliability in the number of correct one-shot decisions. But the immediate problem with any claim of this sort is that we have to have some independent standard or "test" of correctness in mind to form such an expectation. In other words, we have to have some substantive idea of what the right answers to all the likely decision problems *are* before we can assert that any particular procedure will make more right than wrong judgments. But given that we are trying to argue that democracy is the optimal procedure in this sense, we cannot predicate our argument on the assumption that we already know what's right (besides—otherwise democracy would presumably not actually be optimal). Truth-tracking arguments must therefore find some way around the following *Uncertainty Condition*:

To the extent that we cannot independently assert what the truth status of a given P is, we cannot ascertain whether any given judgment of P's truth or falsehood is correct; therefore we also cannot identify whether a given mechanism has a high or low likelihood of giving the correct verdict on P.

Now the analogy between truth-tracking and the jury becomes even clearer. In both instances we are interested in the rightness and wrongness of individual decisions, and we judge procedures according to the ratio of right and wrong decisions they make. A jury has the task of convicting those who are guilty, and acquitting those who are innocent.<sup>156</sup> However, we cannot really measure the performance of a jury unless we know independently who is guilty and who is not.<sup>157</sup>

---

<sup>156</sup> I am indebted to Jon Elster for suggesting this illustrative example.

<sup>157</sup> See also Elster, *Securities against Misrule*.



We could compare different juries in lab settings, using constructed cases where the answer is known ahead of time; but the external validity of such tests depends on whether constructed cases are relevantly like real cases faced by a jury—and to ascertain this we would again need independent access to the true facts about guilt and innocence. The same holds for the political sphere as well. We are looking for a reliable procedure to ascertain what is right or wrong precisely because we do not know what is right and wrong. If we already had a reliable way to ascertain that, why do we not just use that procedure?

Now there is an obvious reply: It might not be possible to ascertain the correctness of a jury verdict *ex post* because we lack independent access to the truth both before and after the trial. We can, however, for instance, potentially verify the correctness of other kinds of judgment—such as predictions—at least after the fact. So perhaps the *past* reliability of certain procedures *can* be ascertained. Accordingly, we may be able to infer a procedure's expected reliability from an extrapolation of its past performance.

This fundamental assumption, for instance, underlies Dietrich's argument it should be unlikely that over time, any procedure would consistently sustain a reliability below chance: after a while people will realize that so many of their decisions turn out wrong that they could actually substantially improve their reliability just by randomizing, or simply by always doing the opposite of what they think.<sup>158</sup> Consistently to perform worse than chance, therefore, is irrational—if I find myself with such a terrible record, there is a simple strategy available: just reverse all my decisions. So the analogy of

---

<sup>158</sup> Franz Dietrich, "The Premises of Condorcet's Jury Theorem Are Not Simultaneously Justified," *Episteme* 5.01 (2008): 56–73.

political procedures to the jury might just be misleading. From the past record of a jury we might not be able to glean anything about its reliability; but political procedures may be very different. So can we ground a reasonable expectation of future reliability in *past* reliability of particular political institutions?

However, this would not only require sufficient evidence to establish such a record, but also a problem context that is consistent over time, and the timely manifestation of measurable results. For a narrow set of well-defined decision problems in data-rich contexts, where immediate feedback is available, it would be relatively easy to form such an expectation. For instance, we can form a reasonable estimate of the reliability of the weather report based on past prediction success. This is because the universe of possible decision problems for the weather report is very narrow, and so the inference from past success to future reliability is more plausible. Furthermore, the weather report's prediction is conclusively (and easily) confirmed or refuted in a timely fashion—by looking out of the window the next day.<sup>159</sup> However, political problems are almost by definition irregular situations, where the usual existing solution strategies do not necessarily apply. Some elements of policy-making, such as annual budget projections, might share the problem structure of the weather report, but a broader view across policy fields suggests matters are more fluid and inconstant.

Nevertheless, I acknowledge that in principle it might be possible to establish the historical reliability of democratic procedures in this way, and also to extrapolate into the

---

<sup>159</sup> However, even here there is a complication: we cannot easily confirm or disconfirm whether the Weather Report's prediction of "a 30 per cent chance of rain tomorrow" has come true or not. Even more problematically, such a statement can be *true* only if we believe in the existence of objective probabilities. For that reason, weather forecasters should be more careful to state what they refer to is only the statistical regularity: "on 30 per cent of days following days with the same weather pattern as today, it has rained."

future. There is some evidence that points in this direction—consider Sen’s claim that there has never been a famine in a democratic state,<sup>160</sup> or the frequently-cited theory of the “democratic peace”<sup>161</sup>. However, in many contexts it is rather more unclear that democratic systems do any better than, for instance, professional forecasters or even market mechanisms. Just some anecdotal evidence suggest that democracies can (under certain conditions) commit even Estlund’s primary bads: Democracies suffer or have suffered from extreme poverty and malnutrition (e.g. India<sup>162</sup>), high homicide rates and/or high rates of incarceration (e.g. Brazil, Honduras, USA),<sup>163</sup> economic collapse (e.g. Weimar Republic, Greece since 2011) and Civil War (USA, Northern Ireland, Yugoslavia), not to speak of decisions that impose extreme costs mainly on *other* countries (consider British rule in India). Now one may analyze the relative likelihood of any of these primary bads, and it may turn out that on average they are lower in democratic governments. However, there is a second problem: it might be very difficult to disentangle the effect of decisions being made *democratically* from the effect of other factors present, such as an efficient bureaucracy, relative wealth, liberal values, the rule of law, or the form of economic organization. So the good decisions democracies make may not have anything to do with the fact that they are taken *democratically*, and everything with contingent factors. Therefore, it remains questionable to claim epistemic reliability for

---

<sup>160</sup> Amartya Sen, “Democracy as a Universal Value,” *Journal of Democracy* 10.3 (1999): 3–17.

<sup>161</sup> Elizabeth Anderson suggests that these empirical facts should be taken to ground an argument for democracy’s epistemic reliability. See Anderson, “An Epistemic Defense of Democracy: David Estlund’s Democratic Authority.”

<sup>162</sup> [http://www.unicef.org/infobycountry/india\\_background.html](http://www.unicef.org/infobycountry/india_background.html)

<sup>163</sup> By contrast, some non-democracies have extremely low rates of homicide (e.g. Singapore, United Arab Emirates, Morocco). See <http://www.unodc.org/unodc/en/data-and-analysis/homicide.html> (accessed April 2013)

democracy on this historical basis without giving a systematic reason why we should expect democracy in particular to decide correctly more often than not.

Finally, the demands of political justification require that democracy should be better not only than *actually existing* alternatives, but better than *potential* alternatives as well. If the only possible alternatives to democracy would be Soviet-style Leninist one-party rule and authoritarian strong-man dictatorship as they existed throughout history, the job of justifying democracy *compared to those examples* would be an easy one. However, we must also consider more “reasonable-seeming” alternatives, such as a democracy with a marginally restricted franchise (where extremist positions are banned, for instance), various unelected technocratic or bureaucratic elements within (such as an independent Central Bank)<sup>164</sup>. How democratic *decision-mechanisms* compare to those alternatives seems a more difficult question. This is not to say that it is not possible to make such a case, but that it *is* necessary to make the case.

However, perhaps the conclusion I have drawn is too strong. After all, we only have to form a reasonable expectation about the average one-shot reliability, and we might be able to do so on purely formal grounds. This idea is behind the recently popular notion of the “Wisdom of the Crowds.”<sup>165</sup> These types of arguments point out that some formal features of democratically organized large groups imply that the decisions made by these groups are more likely to be correct than the decisions made by any of the individual members or subsets of members. The reliability of the crowd, as it were, is

---

<sup>164</sup> See for instance, Frank Vibert, *The Rise of the Unelected : Democracy and the New Separation of Powers* (Cambridge: Cambridge University Press, 2007).

<sup>165</sup> Surowiecki, *The Wisdom of Crowds*.

based on purely formal features of how the crowd is organized, not on any substantive independent understanding of the truth.

This sounds relatively straightforward: we would have to show that these formal conditions obtain, regardless of what turns out to be the right answer. However, there is a problem. Let us call this the *Uncertainty Preservation Objection*:

In so far as the formal conditions under which a mechanism is expected to have a high truth-tracking ratio are defined at least partially with reference to the truth that is supposed to be tracked, the uncertainty about whether the conditions hold in a given situation preserves the initial uncertainty about the truth.

This is again a simple enough point: if the formal mechanism proposed functions only under specific conditions, and if I do not know whether these conditions obtain in the expected circumstances, then I don't know if the mechanism will indeed be reliable. This is especially the case if I need to know what the truth is in order to ascertain whether the conditions hold.

I will discuss arguments for democracy's truth-tracking wisdom in some more detail in chapter 5, where I discuss *particular* forms of organizing democratic decision-making. However, the point ought to be clear in any case: the cogency of a formal (axiomatic) argument that ascribes a certain level of *truth-tracking* to a given socio-epistemic procedure depends on the cogency of the assumptions of the model underlying the procedure. The CJT, for instance, depends on two conditions: that all judgments that enter the majority vote are statistically independent, and that the average judgment has a likelihood of being right that is higher than chance. So even if we do not know what the right answer is, so it seems, we can trust in a majority vote to get it right if we trust that the assumptions hold.

However, the obvious problem is of course that *unless we know what the right answer is*, we do not know whether the competence condition holds. The same point that applies to the historical evaluation of reliability discussed above also holds here. Now one may think that an assumption of being better than chance *on average* is a pretty low threshold for this key assumption. However, we have to be clear that in order to make such an assumption we have to rule out, among other things, that the average voter is subject to misleading information or common heuristics and cognitive biases. Now whether or not this is actually a cogent assumption, we cannot completely *rule out* whether a set of information is misleading unless we know what true information (i.e. non-misleading information) would be. The same goes for cognitive biases insofar as they pertain to this problem: we can only judge whether people are subject to say, distorting framing effects, if we have an idea of what it is that is being distorted.<sup>166</sup>

So the point I am making is *not* that we have good reason to believe that the conditions of the CJT is regularly violated: rather, we just don't know with particular high certainty. But pragmatic robust instrumentalism demands that we can make such a claim with the requisite certainty. To put it differently, assume that I try to justify the authority of a particular democratic decision to you (let's assume regarding a decision with which you disagree), and I attempt to use the CJT. I tell you that a decision by a large group, taken by majority vote, is likely to be the *right one*—if the conditions are fulfilled. Let's then assume that you doubt that the competence condition is fulfilled: since you do not agree with the decision, you presumably do not think that people are—on average—right about

---

<sup>166</sup> Of course we may think that in general, any information that is presented where there is self-interest in play is liable to be distorted, but that itself may be hard to ascertain. For a highly pertinent discussion of democratic theory and framing effects, see Jamie Terence Kelly, *Framing democracy : a behavioral approach to democratic theory* (Princeton: Princeton University Press, 2012).

this. Maybe you actually have a good reason for this: e.g. you think that people are influenced by distorting media. So what reason could I give you to change your mind on the grounds of observing the outcome of a democratic decision?<sup>167</sup> In particular, can I give you a reason that is *not* premised on an understanding of what I think the right decision actually is? How can I counter your claim that the picture presented by the media is distorted, in a way that you have to accept?

The crucial idea is that whether or not the competence condition holds, a reasonable person may well disagree with my assumption that it does. This is a general problem with the *application* of axiomatic models to argue for the epistemic virtue of certain real procedures. This includes modifications of the CJT<sup>168</sup> as well as diversity-based theorems such as Page's *Diversity-Trumps-Ability* Theorem.<sup>169</sup> The conclusions of these theorems of course always follow from the assumptions by definition: models cannot be “wrong” per se. So the important question is whether we can confidently make a case that the assumptions hold. This again, as I argue, depends on what we think the right solution is. If we take disagreement to have a normative role, citing a theorem of this kind is not sufficient to ground political legitimacy. I will for now postpone further discussion of these “wisdom-of-the-crowds” mechanisms to chapter 5, where I discuss different concrete models of democracy.

This section was only concerned to establish that the *standard* way of arguing for the epistemic reliability of democracy—what I have identified as the *truth-tracking* or *juror*

---

<sup>167</sup> A similar argument is made by Ingham, “Disagreement and Epistemic Arguments for Democracy.”

<sup>168</sup> Dietrich and Spiekermann, “Epistemic Democracy with Defensible Premises.”

<sup>169</sup> Page, *The Difference*.

model, is insufficient to ground the legitimacy of democratic authority in the robust way we need. As per the PRI principle above, while citizens may have an interest in the existence of a procedure that resolves political problems, we cannot claim that democracy fulfils this function by having a high ratio of hits over misses in its decision-making. Is there a way to think differently about reliable problem-solving?

## 5. Scientists, Not Jurors

Now the previous section reached what seems like an odd and implausibly radical conclusion: that in general there should be no way to judge the epistemic quality of different procedures unless we know what we are trying to discover. *Normally* we do not know what it is that we are trying to discover, yet in everyday life we make such judgments all the time. Even though we don't know what future research in biochemistry will discover, we think a team of scientists is better suited epistemically to the task than a team of kindergartners. Even though we don't know how to make the best possible bread (we don't have a picture of the ideal bread in our minds), we can confidently say that a baker will have a higher epistemic reliability at making good bread than the author of this chapter. And these judgments seem to be both reasonable and reliable. If my standpoint denies the validity of such claims it seems hopelessly over-skeptical. So what is going on?

The problem is that the measure of *reliability* employed by the Juror Model is *too demanding*: to maximize the ratio of correct over false decisions or to "track the truth" cannot be the standard by which we measure political procedures. As I have argued, we



cannot reliably say which procedure is more or less reliable in this sense unless we know already which decisions are right and which are false.<sup>170</sup>

However, outside the courtroom, we actually make these kinds of judgments of epistemic reliability quite *differently*. A scientist, engineer, or baker is not assessed with respect to her likelihood of finding “the truth” in the sense of a true description of the world, the “right” way to design an airplane wing, or the “perfect” bread. Rather, what we are looking for is that she will find *pragmatic solutions* or *progressive improvements*. These are not necessarily *true*, except in the *pragmatic* sense of the word: they might not be optimal, might not even be approximations to any objective optimum, and need not to be expected to be permanent. Here, progress is defined with respect to the problem to be solved: to explain observed phenomena better than previous theories, to make airplanes fly more efficiently and safely, and to make tastier bread. Replacing the Ptolemaic system of astronomy with the Copernican is a progressive solution when it comes to explaining the phases of Venus; but the Copernican system is still inadequate when you want to fly to the moon, let alone correct in some objective sense.<sup>171</sup> As J.S. Mill puts it very nicely in a proto-pragmatist passage: “Even progress, which ought to superadd, for the most part only substitutes, one partial and incomplete truth for another; improvement consisting chiefly in this, that the new fragment of truth is more wanted, more adapted to the needs of the time, than that which it replaces.”<sup>172</sup>

---

<sup>170</sup> Note also that in contrast to baking or chemistry, where we have a general idea about this, we have absolutely no idea about the number of wrongful convictions, and much less about the number of wrongful acquittals.

<sup>171</sup> I am indebted to William McAllister for suggesting this useful example.

<sup>172</sup> Mill, “On Liberty” 114.

So, the resolution of a problem may not require us to know the “truth” in a transcendental sense. In order to address a practical issue in the political realm, we need not figure out what the optimum or “ideal” decision may be. This is a problem-focused idea of what the epistemic function of a political procedure is. The key point is that on this view, political activity is not a branch of moral philosophy, as it were—trying to figure out what is right, and then attempting to realize it. Rather, political activity is trying to figure out what is *wrong*, and trying to overcome this.<sup>173</sup>

Of course the fact certain actions or decisions *are* progressive improvements fundamentally depend on certain true or false *facts* about them. In other words, whether or not a certain decision *is* in fact a solution. People will like the bread in virtue of its molecular composition, and the new airplane wing will work better because of certain underlying physical principles. Now, of course this *seems* to leave us with the same problem again, just at a different level. How can we be sure that our procedure will actually reliably solve political problems? Even if we don’t expect truth-tracking, should we not expect something like solution-tracking?

However, even if the effectiveness of potential political solutions depends on facts, remember that in evaluating the reliability of an engineer and a baker, we are in the domain of *normative epistemology*, that is, of theories of what the goal of an epistemic inquiry should be. When we understand the task of the political system in this way, the relevant epistemic goal here is *not* to maximize the ratio true over false beliefs, or correct over false decisions. We are not primarily interested in a high average one-shot likelihood that a given policy will actually solve a problem.

---

<sup>173</sup> Sen, *The Idea of Justice*; Wiens, “Prescribing Institutions Without Ideal Theory.”

As it were, if the long term, diachronic goal of policy-making is the resolution of political problems, what matters is whether those problems will eventually get resolved, not how many false attempts are made (within reasonable limits). Furthermore, with those professions, we do not even expect high individual reliability from the baker or scientist, as long as she employs the right *adaptive strategy* to deal with the results of past decisions.

I propose therefore that in the political context we should see epistemic reliability *more like science, engineering, or baking*, and less like *attempting to get as many verdicts right as possible*. Politics, in this view, should be seen less like a method of delivering true judgments on justice, morality, or factual questions. This process may have the side-effect that we find out the truth about certain facts (those by virtue of which our policies resolve our problems), but this is not the main goal, and crucially, it is not the standard by which political decision-making should be measured.

### *Pragmatic Reliability*

Now, viewing politics in this way has some useful consequences. Firstly, I believe this more adequately reflects what politicians are actually doing. Secondly, however, we can understand better how we *ascribe reliability* to different epistemic procedures, namely by looking at their *methodological* features. In particular, it gives us a handle on how to see democracy's key epistemic advantage, and how to give an epistemic justification of democracy while avoiding many of the critiques leveled against it. The vital point is this: we make judgments of *long-term epistemic reliability* without claiming a particularly high *truth-tracking ratio*. Science, to return to this example, is an inherently diachronic, continuous

endeavor. Accordingly, the epistemic value of a functioning scientific procedure does not lie in the likelihood that any of the hypotheses turn out right—but that at any given time, we have reason to trust that the currently prevalent hypotheses and theories represent the currently best attempt at explaining the evidence. And whether a procedure is reliable in this sense, we can answer by looking at its methodology.

Consider a straightforward example: We do not know what the perfect bread looks like, and we do not have a simple linear scale of the quality of bread (e.g. “the crustier the better”). So then when do we think a particular baker will improve the quality of bread? It is not that the next bread she tries out is *more likely* to be the perfect bread. A good baker might well have a terrible ratio of right versus false decisions: if she is an experimental type, she might only make one good type of bread for every fifty she tries. But that is fine: we assess the epistemic adequacy of a baker by the *methods* she employs going about the epistemic task. The progressive baker experiments and adapts to the outcome of the baking. The bad baker either fails to experiment at all (never tries out new bread) is unwilling to adapt to the results of experimentation (keeps on making a variety that includes many types of bad bread), or lacks the ability to enable successful adaptation (cannot tell which bread is good or not).<sup>174</sup> Of course *afterwards* we do tend to rationalize the story of a successful baker: she must have had the special *je-ne-sais-quoi* that gave her a high likelihood of discovering the good bread. But that does not mean that we could have told that same story *ex ante*.<sup>175</sup> From this perspective many of our common-or-garden

---

<sup>174</sup> In an ideal marketplace of course this baker would eventually be removed by losing business. Alas, the market is imperfect and bad bread persists.

<sup>175</sup> this tendency of human behaviour of ex post rationalization and the focus on patterns is admirably discussed in Watts, *Everything Is Obvious: Once You Know the Answer*.

epistemic judgments also make more sense: the biochemists have a higher reliability than the children because they have the skill to *conduct experiments* and they can assess whether an experiment has worked or not, not because they have an inherent tendency to discover the true answer *on a given question*.

Similarly, when we judge *ex ante* the diachronic epistemic reliability of scientists we do not try to estimate their expected ratio of making right over wrong judgments, but we assess the way in which they approach problems: according to some version of the scientific method, which involves constantly subjecting even apparently secure hypotheses to potential falsification and reinterpretation. What makes one a good scientist is not a certain level of *ex ante* truth-tracking capacity. That kind of thing cannot be inferred ahead of time, since we do not know what the right answers are to scientific questions (this might also be the reason Nobel Prizes are not awarded *ex ante*), but a willingness to revise or discard hypotheses and even full-blown theories in the face of evidence, and to continuously subject apparently settled views to experimental challenges. And even if the evidence should conform to the theory, the truth of the theory can be asserted only provisionally: it might be overturned by different—or better—evidence, and made obsolete by further progress.

From a longer-term point of view, therefore, the value of the scientific, experimental “strategy” does not depend on whether any given individual decision is more likely to be true than not, even if the methods are employed correctly. In the same way we cannot just take snapshot of the current state of some scientific field and argue that it is particularly likely to be *correct*. What we can hope for is a high likelihood, over time, to come up with progressive solutions: coming up with new ideas, eliminating those

that are refuted, superseded, or no longer adequate for their purpose, and to corroborate those that are not, altogether adding to the record of knowledge, without the assumption that at any given point all (or any!) of our beliefs are true in the objective sense.<sup>176</sup>

More generally (e.g. across most fields of inquiry), in circumstances of deep uncertainty we should think of epistemic advantage in that way: given such uncertainty, we might not be able to say which *decision* is objectively correct, but the adaptive *strategy* shown by the good baker seems the adequate response. This criterion of epistemic success is *pragmatic*—the fundamental epistemic concept is not that of *truth*, strictly speaking, but that of *progress*, where progress is relative to a specific practical problem context. Inventing a new type of bread can be progressive when the new product addresses the “problem” of finding a better bread, but it is hard to argue that it is better according to some arcane objective scale of goodness.

## 6. A Pragmatic Model of Politics

Once we look more closely at the political sphere, we realize that political activity actually more resembles this problem-solving than things like jury duty. Political decision procedures are not usually in the business of passing truth verdicts on certain propositions. What they are typically doing is designing and implementing rules, orders, and programs that resolve particular problems; usually problems arising from a clash between different value systems and social phenomena. Democratic procedures of course

---

<sup>176</sup> This point shares much with Elizabeth Anderson’s sketch of a pragmatist model of democracy, although my emphasis is slightly different. See Elizabeth Anderson, ‘The Epistemology of Democracy’, *Episteme* 3, 1-2 (2006).

have to judge the expected quality of different alternatives—say, alternative policy proposals—but the judgment is on the context-specific merits of the proposals rather than purely their truth. Philip Kitcher has argued that this is the function not only of politics, but also of the practice of ethics generally. As he puts it: “ethical progress is prior to ethical truth, and truth is what you get by making progressive steps (truth is attained in the limit of progressive transitions; truth “happens to an idea”)”<sup>177</sup>

In the pragmatist view, therefore, we can see *progress as functional refinement*: a change towards a new policy, ethical rule or norm, or theory about the world constitutes progress if it fulfills its function better than the previous policy. In a political sense, therefore, progress is of course defined not globally, but only with respect to social *problems*: if policy P solves problem X, then it makes progress with respect to X. P is *correct* for all intents and purposes, even though it might not be the only possible solution and therefore not the only step that would be progressive. In the context of justification: “If and when we need a notion of ethical justification, it is easily found; people are justified when their decisions are generated by processes likely to yield progressive changes. Reliability in the production of ethical truth gives way to reliability in the genesis of progressive transitions.”<sup>178</sup> In the same way we could justify political authority: if democracy is likely to reliably generate progressive transitions, it can be justified.

A progressive transition does not even have to be optimal. Indeed, policy P might give rise to another serious problem Y, which requires another attempt at problem-

---

<sup>177</sup> Kitcher, *The Ethical Project*; 210. The quoted phrase is of course from William James. The full quotation is, “Truth *happens* to an idea. It *becomes* true, is *made* true by events.” See William James, *Pragmatism, Great Books in Philosophy* (Amherst, MA: Prometheus Books, 1907), 89.

<sup>178</sup> Kitcher, *The Ethical Project* 212-213.

solving, and progressive step R, in an unending chain of problem-solving. We might never find the conclusive set of ideal policies, and therefore political activity, in so far as politics is applied problem-solving, might never end. Progress, in the Deweyan sense is not progress *toward* something; it is just open-ended progress, increasingly sophisticated problem-solving, full stop. It is important to note that progress in this sense is not *global*, i.e. progressive “in the grand scheme of things.” In a pragmatic normative epistemology, we do not need such a “view from nowhere” perspective from which to judge the global progressiveness or epistemic adequacy of a given political or other choice.<sup>179</sup>

Note once more that the switch to a pragmatic understanding of political reliability does not necessarily entail a commitment to a pragmatic understanding of the “truth.”<sup>180</sup> It does, however, entail a commitment to the idea that ascertaining the truth-status of the assumptions underlying political decisions is not the primary *goal of inquiry*. Of course, policies might work *in virtue of* these truths about factual and normative questions. But this does not affect how we should understand the epistemic role of *political activity*: reliability in the realm of politics does not consist in the maximum number of *true* policies at any given point. Instead, a procedure is reliable when at any given point we can expect it:

*Either* (i) to represent our best current attempt to resolve all currently relevant political problems, given current circumstances,

*Or*, (ii) if there are unresolved political problems, to have an adequate adaptive methodological strategy to resolve them.

---

<sup>179</sup> Ibid.

<sup>180</sup> See also the discussion on pp. 137-8 above



*And* (iii) to have an adequate adaptive methodological strategy to resolve future political problems should they occur.

On the pragmatic understanding, this grounds the Reliability Claim that forms the basis of the PRI principle. The key consequence is this: as it turns out, we can form this expectation without reference to the truth about problems: if a procedure is experimental and adaptive, we can reasonably expect it to be reliable in this sense.

Is there something lost in this picture? Well, one might think that this is an unsatisfactory understanding of what politics is about, and that political authority—if it can ever be justified—*should* be concerned with figuring out the fundamental *truths* about justice and rights, and the objectively best way to realize them: that it should approximate the perfect judge; and that politics should really *track the truth* in a fairly substantive sense. I cannot not address this here, but I acknowledge that that this is a possible understanding of the task of politics. However, if we see it that way, we are saddled once again with the problems of ascertaining reliability that has been the subject of sections 4 and 5—and unless we can overcome them somehow, we might have to abandon the idea of an epistemic justification in the first place.<sup>181</sup>

## 7. Conclusion

So now we have the basic structure of the argument. We required a justification for democratic political authority that takes account of the normative importance of *outcomes*

---

<sup>181</sup> I will also briefly return to this point in Chapter 5 below.

while acknowledging the depth of disagreement and uncertainty about how to produce good outcomes, and even about what they are. This chapter has argued first, that even under conditions of disagreement about the ultimate ends of politics, reasonable political agents have an interest in the existence of a political decision procedure (a “state”) that *they* can expect to resolve common political problems. This was termed the principle of Pragmatic Robust Instrumentalism (PRI). Even under conditions of disagreement we can speak of better and worse decisions—deep disagreement indicates *uncertainty*, and the adequate response to uncertainty is to proceed with inquiry rather than to give up on the possibility of having any standards. However, the problem with giving such a reason was that it must be acceptable from all reasonable points of view: a full justification of democracy on those terms must be *robust*. This means that any claim to democracy’s capacity with respect to this problem-solving function may not be predicated on specific understandings of the truth (those that may be reasonable rejected, given the uncertainty).

The latter half of the chapter argued that this means that the typical epistemic justifications of democracy do not actually work. The reason is that they tend to conceptualize the epistemic *task* of democracy as a high likelihood, on a given decision, to get it right. Instead I have advocated to understand reliability diachronically and dynamically as the adequate methodological approach to resolve problems of uncertainty over time. This may involve getting many decisions initially wrong—but this does not matter if these are continually subjected to experimental re-evaluation. On this basis I have argued that we should interpret “reliability” in the sense of the PRI as *pragmatic problem-solving capacity*. Accordingly, in order to give a robust justification of democracy, we should focus on democracy’s ability to actually employ such an experimental decision-

making strategy. This shifts our focus on to the functional features of democracy that are relevant for this: bringing problems onto the agenda, creating a diversity in potential solutions, and particularly also the capacity to enable adaptation through feedback and monitoring. However, before I flesh out the particularly *democratic* part of this argument in the last two chapters, I move on to discuss some objections. Chapter 4 is concerned with the question of whether an experimental strategy of making binding, coercively-enforced political decisions can at all be justified.

---

## Can Experimental Political Authority Be Justified?

### 1. Introduction

I begin this chapter by contrasting quotations from two politicians of—so it seems—very different temperaments. First, Franklin Delano Roosevelt, from his 1932 Commencement Address at Oglethorpe University:

“The country needs and, unless I mistake its temper, the country demands bold, persistent experimentation. It is common sense to take a method and try it: If it fails, admit it frankly and try another. But above all, try something. The millions who are in want will not stand by silently forever while the things to satisfy their needs are within easy reach. We need enthusiasm, imagination and the ability to face facts, even unpleasant ones, bravely. We need to correct, by drastic means if necessary, the faults in our economic system from which we now suffer. We need the courage of the young. Yours is not the task of making your way in the world, but the task of remaking the world which you will find before you. May every one of us be granted the courage, the faith and the vision to give the best that is in us to that remaking!” (Franklin Delano Roosevelt, *Oglethorpe University Commencement Address* (1932))

In contrast, consider Konrad Adenauer’s concise campaign slogan for his 1957 re-election campaign for the German chancellorship. This was during the height of the Cold War, of course, and the political circumstances in Germany were not free of uncertainty

and risk by any means. Nevertheless, it is widely considered one of the most effective campaign slogans in postwar German political history:

“Keine Experimente!” (No experiments!)

This basic difference in attitude towards experimentation in the political sphere is of course partly due just to historical circumstances, but it may also reflect a deeper difference in political ideals: progressivism vs. conservatism perhaps, but importantly also a fundamental difference in attitude towards risk and the virtues of confronting uncertain situations with experimentation. That Adenauer’s re-election campaign could be based merely on the promise *not* to experiment seems remarkable, and it suggests that the appeal of experimental policy-making, though perhaps initially intuitively plausible, is not as universal as it seems. Since on the other hand the last chapters have outlined a theory of democratic legitimacy that focuses on precisely this idea of efficient experimental policy-making, this difference has to be explored further. The chapter therefore takes Roosevelt’s side in this debate, and tries to defend the view that “bold, persistent experimentation” can be robustly justified—and by implication that a democratic system that is essentially based on an experimental form of policy-making can command legitimate authority.

The basic normative assumption that this chapter examines—and with which the last chapter ended—is that we should understand political decisions as *experimental interventions* (as opposed to, say, *verdicts* or *statements of fact*), and that their quality ought to be evaluated accordingly. The activity of policymaking in this model resembles *engineering* more than *judicial proceedings*. As I have argued above, understanding politics in this way

has a number of important implications. First, it means that we get a different view of when a policy process is working well and when it is not. Second, we get a new perspective on what it is about *democratic procedures* that is particularly valuable.<sup>182</sup> What is the actual value of deliberation, participation, representation or the other mechanisms of democratic politics? Third, it also requires us to re-evaluate the *justification* of political authority. If we think that politics is mostly “just” more-or-less educated stabbing in the dark, how can we possibly treat laws as imbued with normative authority? Can experimenting ever be the right thing to do, politically? These are the questions I am dealing with in this chapter. My general answer is: if the principle of pragmatic robust instrumentalism that has been introduced in the last chapter can be justified, then on this basis the authority of experimental policy-making can also be justified.

Recall that the motivation for the idea that policy-making should be more *experimental* and/or *adaptive* results from the idea that it is in every reasonable person’s interest that there should be a stable political decision-making authority that effectively resolves political problems and conflicts that arise because of deep uncertainty (on a normative as well as a factual level). Now this idea that the adequate response to uncertainty is an *experimental* strategy of inquiry or problem-solving is not unique to political philosophy.

This general idea has recently received a lot of attention, from several different areas: the social sciences (especially decision theory and complexity theory), epistemology and ethics, as well as actual policy-making. The general background to this seems to be the realization that the social world is deeply complex and diverse, composed of adaptive

---

<sup>182</sup> This point will be expanded in the next section.

individuals interacting and shaping their beliefs and expectation in response to these interactions. In other words, that the social world is a *complex adaptive system*. Now, the problem with such a system is that its macro-behavior depends on the aggregate of many interrelated micro-level decisions, which themselves are crucial environmental factors determining other micro-level individual decisions.<sup>183</sup>

Hence the full consequences of any particular policy—and indeed the behaviour of complex systems such as the economy as a whole—are *ex ante* not completely predictable. If this is true—and the (lack of) predictive success with respect to those systems suggests that it might be—this has some important consequences on how we should understand politics.

But before we get to these arguments, we should think about what it means to have an experimental policy-making strategy. The first thing to note is that there is no *one correct* experimental methodology: “There is no perfect or true experiment. The appropriate experimental design depends on the research question just as is the case with observational data. In fact, the variety of possible experimental designs and treatments is just as wide as it is with observational data, and in some ways there is a greater range of possibilities through experiments than is possible with observational data.”<sup>184</sup>

Uncertainty demands certain things from our policy-making. First, policy should be designed with the possibility in mind—or even the expectation—that it could go *wrong*,

---

<sup>183</sup> Sandra D Mitchell, *Unsimple Truths : Science, Complexity, and Policy* (Chicago: University of Chicago Press, 2009); Scott E Page, *Diversity and complexity*, *Primer in Complex Systems* (Princeton, NJ: Princeton University Press, 2011).

<sup>184</sup> Rebecca B Morton and Kenneth C Williams, “Experimentation in Political Science,” in *The Oxford Handbook of Political Methodology*, ed. Janet Box-Steffensmeier, David Collier, and Henry Brady (Oxford: Oxford University Press, 2006), 5.

i.e. that it will not actually fulfill the intended function, or that it has unpredicted consequences that lead to equally significant problems somewhere else. Accordingly, it is essential that the consequences of policy are *monitored* and that policies keep on being revised or even revoked in response to the outcomes. Second, in so far as policy creates autonomous complex adaptive systems through its regulation, such as the financial markets or the labor “market,” there is a need for these systems to be *robust* and *adaptive*. Robustness refers to the stability of a system in response to external shocks. A robust system maintains its (optimal or satisfactory) functionality under a wide range of possible scenarios. If the functioning of, say, the financial markets depends on the accuracy of a single risk-estimation formula, it is not robust, but rather *fragile*. Adaptiveness, that is, the capacity to change the functional structure of the system in response to emerging problem, might contribute to robustness.

Third, however, policy-making ought to be structured more like an *experiment*: that means policy should have clearly articulated goals, but should be deliberately provisional, subject to clear, ascertainable standards and implemented in a controlled way; and should be subject to re-evaluation in response to the observed outcomes. These are the three functional elements of an experimental strategy—even though there are large differences in how these are concretely realized in particular instances.

## 2. Complexity, Uncertainty, Difficulty

What accounts for the growing popularity of experimental approaches in social-scientific contexts, including in actual policy-making situations? The general background to this



seems to be the realization that the social world is deeply complex and diverse, composed of adaptive individuals interacting and shaping their beliefs and expectation in response to these interactions. In other words, that the social world is a *complex adaptive system*. Now, the problem with such a system is that its macro-behavior depends on the aggregate of many interrelated micro-level decisions, which themselves are crucial environmental factors determining other micro-level individual decisions.<sup>185</sup>

Hence the full consequences of any particular policy – and indeed the behaviour of complex systems such as the economy as a whole – are *ex ante* not completely predictable. If this is true – and the (lack of) predictive success with respect to those systems suggests that it might be – this has some important consequences on how we should understand politics.<sup>186</sup>

The application of an experimental mode of policy-making to decision-making under uncertainty has its origin in complexity research, and particularly the analysis of complex adaptive systems. One implication of this research is that the behaviour of those systems, especially in response to external stimuli, can be close to unpredictable.<sup>187</sup> The consequence of this unpredictability is that the main alternative to experimental policy-making—let us call this *optimization*—is unavailable. As already briefly mentioned above, this approach to decision-making under uncertainty relies on a two-step model: first the

---

<sup>185</sup> See for instance Mitchell, *Unsimple Truths : Science, Complexity, and Policy*; Page, *Diversity and complexity*; Page, “Uncertainty, Difficulty, and Complexity.”

<sup>186</sup> Of course there are areas in which prediction is possible, and indeed where our ability to do so is improving. This includes contexts where vast amounts of data can be collected (“big data”), and/or where fundamental causal relationships are relatively well understood. But as Nate Silver, one of the most successful political forecasters of the last years confirms, those areas are few and far between. See Nate Silver, *The Signal and the Noise* (London: Allen Lane, 2012)

<sup>187</sup> Mitchell, *Unsimple Truths : Science, Complexity, and Policy*; Page, *Diversity and complexity*.

estimation of the respective probabilities of expected consequences, i.e. the prediction, and then the maximization of the expected payoff, given those estimates. Brute uncertainty implies that step 1 is not available to us: but how may we understand that?

The idea that humans are inherently bad at predicting consequences of purposive actions is a common and long-established theme in psychological and popular science literatures. For reasons of bounded rationality and cognitive biases,<sup>188</sup> because of overconfidence in one's predictive abilities,<sup>189</sup> or because of a tendency to see patterns in historical developments and ascribe cause-and-effect relations to contingent developments,<sup>190</sup> individuals tend to be not only bad at predicting, but also good at rationalizing the failures of their predictions and therefore at refusing to adapt to these failures. This already suggests that predict-and-act optimization models might not be the only adequate standard: we might think that a good procedure also tends to realistically assess its own likelihood of failure. Nevertheless, we might think that these individual cognitive shortcomings could be overcome—and specifically that collective mechanisms (such as deliberation) could serve to ameliorate these biases.

However, complex systems in general often *in principle* exhibit unpredictability of this kind. I shall here just mention three features of complex systems that lead to unpredictability: emergence, chaos, and critical states. The term “emergence” describes features of mass behaviour that are not reducible (in explanatory terms) to individual

---

<sup>188</sup> Daniel Kahneman, Paul Slovic, and Amos Tversky, *Judgment Under Uncertainty : Heuristics and Biases* (Cambridge: Cambridge University Press, 1982); Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011).

<sup>189</sup> Philip E Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, NJ: Princeton University Press, 2005); see also the concept of the "planning fallacy" in Kahneman, *Thinking, Fast and Slow*.

<sup>190</sup> Watts, *Everything Is Obvious: Once You Know the Answer*.

behaviour of any of the elements in the mass. So for instance a flock of birds or shoals of fish exhibit emergent behaviour since the precise movements of the whole group depends on minute reactions of the individual birds to where each other bird is going, on two-way interactions between the flock and the individual. The group behaviour “emerges” from the individual behaviours but is not reducible to them (at least not in reasonable amounts of time, and given human cognitive constraints).<sup>191</sup>

As system is chaotic when imperceptible differences in the starting situation have greatly disproportionate effects on outcomes—this concept is called “sensitive dependence on initial conditions.”<sup>192</sup> Chaotic systems are not random—indeed the study of dynamic systems operates from the assumption of full causal determinism. Nevertheless, complex systems, despite being fully determined, can exhibit seemingly random behaviour without external sources of randomness. The famous example of a chaotic system is of course the butterfly in the Amazon basin causing a hurricane across the globe. But in social systems it may be equally the case that imperceptible and contingent differences in the causal history of an event disproportionately influence the outcome. The take-home point here is however that behaviour of complex systems can be seemingly random and therefore unpredictable *without* abandoning the metaphysical assumptions of causal determinism. A social system, for instance, in which everyone is acting perfectly rationally, can still exhibit unpredictable behaviour. Indeed, markets are sometimes cited as examples for systems that are *per se* chaotic.

---

<sup>191</sup> Melanie Mitchell, *Complexity : a Guided Tour* (Oxford England ; New York: Oxford University Press, 2009) 34-44.

<sup>192</sup> Mitchell, *Unsimple Truths : Science, Complexity, and Policy* 20.

Thus, it seems to be entirely unpredictable to where a flock of birds, when startled, will move. Sagalnik, Dodds and Watts' famous Musiclab experiments may be an illustration of emergent behaviour in the social realm as well.<sup>193</sup> In these experiments subjects were offered a choice of free music downloads on a social network site called *Musiclab*. The subject could observe the popularity of the different songs, which were ranked by number of downloads by the other users. Crucially, the researchers randomly assigned users to one of eight different "markets," such that the download statistics they could see were only the statistics of the market they were in. If we believe markets such as this one would be predictable in some sense, we should expect that all eight markets would in the end exhibit the same features—the same songs ought to be on top of the list, and the same songs on the bottom. However, as Watts shows, the different markets developed entirely differently, with different songs topping the lists in the different markets—since consumer choices were so dependent on following what others did in the markets the users as a whole exhibited *emergent* behaviour. Their download choices could not be modelled using their idealized individual choices alone.

Finally, the idea of a *critical state* adds to the unpredictability of events the unpredictability of the *magnitude* of events. Essentially, this means that while in ordinary circumstances we might be able to accurately foresee the behaviour of a given system, large-scale catastrophic events where the whole system shifts into another state are often unpredictable:<sup>194</sup> "It appears that, at many levels, our world is at all times tuned to be on the edge of a sudden, radical change, and that these and other upheavals may all be strictly

---

<sup>193</sup> Matthew J Sagalnik, Peter Sheridan Dodds, and Duncan J Watts, "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market," *Science* 311 (2006): 854–856.

<sup>194</sup> Per Bak and Kan Chan, "Self-Organized Criticality," *Scientific American* 264.1 (1991): 46–53.

unavoidable and unforeseeable, even just moments before they strike.”<sup>195</sup> The idea here is that certain systems organize themselves into a fragile state where small causes can cascade through the system and have catastrophic effects; while in a “normal” state these small causes would have small effects, corresponding to their size. This seems to be a good explanation of how earthquakes and landslides work, and could potentially be applied to social systems as well. The “tipping point” and “cascade” models has been used to model riots and other spontaneous collective actions, but also cascading behavior of other kinds (most famously the serial collapse of the Eastern European Communist regimes in the period of 1989-1991).<sup>196</sup> The key here is that in a critical-state system we can say that with some certainty we can expect that catastrophic events will occur with some statistical regularity, but we cannot accurately predict precisely *when* any individual catastrophic event will occur. Thus, there are constant occurrences of earthquakes that can be quite accurately predicted, but occasionally a small movement of the earth’s crust has wider effects and will lead to a catastrophic earthquake. The occurrence of these events depends on a large number of individual elements being arranged just-so. As such, this means that the usual cause-and-effect explanations are not useful here, since the magnitude of the effect depends crucially on the historical development of the current state. The magnitude of the effect is, as it were, path-dependent, and unpredictable unless we know the whole “path.”

This is similar to the idea of a “black swan,” as Nassim Taleb calls low-probability, high-impact events that are unpredictable. A dynamic social system that is prone to

---

<sup>195</sup> Mark Buchanan, *Ubiquity : Why catastrophes happen* (London: Weidenfeld & Nicolson, 2000) 62.

<sup>196</sup> Mark Granovetter, “Threshold Models of Collective Behaviour,” *American Journal of Sociology* 83.6 (1978): 1420–1443; Timur Kuran, “Now out of Never: The Element of Surprise in the East European Revolution of 1989,” *World Politics* 44.1 (1991): 7–48; Timur Kuran and Cass Sunstein, “Availability cascades and risk regulation,” *Stanford Law Review* 51.4 (1999).

organize itself into a critical state exhibits emergent and unpredictable behaviour that is unpredictable and subject to potentially large-scale change.

I shall not go into much more detail regarding unpredictability of complex systems now. Suffice it to say that there is reason to believe that in many areas of concern to policy-making, it seems that the use of approximations, best estimates, and predictions to stand in for the true probabilities of the possible consequences of actions seems problematic.

*Responding to complexity with experiments*

The consequences of these considerations are that one cannot rationally make optimal decisions under uncertainty based on calculations of expected utility or estimations of objective risks. In other words, *optimization* is not a feasible response when we are faced with such conditions. As Axelrod and Cohen put it, under these circumstances all we can do is *harness* the complexity.<sup>197</sup> By this, they mean *accept* the fact that complexity cannot be entirely controlled, and deliberately set up your system to benefit from positive surprises and exhibit robustness to negative surprises.<sup>198</sup> Experimental policy-making takes this into account. In philosophy, especially in epistemology and ethics, this idea has also gained some traction recently, particularly in the context of a revival of Deweyan and more generally, *pragmatic* theories of knowledge.<sup>199</sup>

---

<sup>197</sup> Robert M Axelrod and Michael D Cohen, *Harnessing complexity : organizational implications of a scientific frontier* (New York: Free Press, 1999).

<sup>198</sup> Taleb calls the capacity to resist and even benefit from deviations "anti-fragility." See Nassim Nicholas Taleb, *Antifragile: Things That Gain from Disorder* (New York: Random House, 2012).

<sup>199</sup> Robert B Talisse, *A Pragmatist Philosophy of Democracy, Routledge Studies in Contemporary Philosophy* (New York, London: Routledge, 2007); Talisse, *Pluralism and Liberal Politics*; Eric MacGilvray, *Reconstructing Public Reason* (Cambridge, MA: Harvard University Press, 2004); Eric MacGilvray, "Experience as Experiment: Some

There is a wide variety of manifestations of the general “pragmatic program” in philosophy; the most salient difference may be between more naturalistically-oriented positions (this strand builds mainly on C.S. Peirce) and more subjectivist positions (in the tradition of William James).<sup>200</sup> Since this dissertation is concerned mainly with a normative theory of democratic legitimacy, and neither with the history of philosophical thought nor the deep questions of epistemology, I will not go too deep here into this discussion. In any case, regardless of this divide, we may identify a number of core aspects that tend to be common to all pragmatist philosophical positions. First, there is the view that the meaning of philosophical concepts is determined by the practical consequences they have. This is coupled with a concern that philosophical *problems*—i.e. calls for inquiry—have to be motivated by *real, practical* doubt. This position, which for instance rejects Cartesian universal skepticism, may be called anti-foundationalism. Philosophical answers, as it were, are concrete resolutions of practical problems. Second, there is the view that since the meaning of knowledge is determined by the real-world effects it has, we have to assume a basic fallibilism about our own knowledge. Furthermore, the “truth” of a concept will only shake out *in practice*, in the concrete application.

We can already see how these basic epistemological commitments motivate an experimental methodology. Pragmatism rejects the capacity of *a priori* theoretical exercises in testing the truth-value of propositions: hence *ex ante* theoretical determinations of

---

Consequences of Pragmatism for Political Theory,” *American Journal of Political Science* 43.2 (1999): 542–565; Kitcher, *The Ethical Project*; Kitcher, *Preludes to Pragmatism*; Cheryl J Misak, “A Culture of Justification: The Pragmatist’s Epistemic Argument for Democracy,” *Episteme* 5.01 (2008): 94–105; Cheryl J Misak, *Truth, Politics, Morality: Pragmatism and Deliberation*, (Routledge, 2000); Misak, *The American Pragmatists*; Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*.

<sup>200</sup> This distinction can be traced throughout the history of 20th century philosophy. See Misak, *The American Pragmatists*.

normative or factual matters is immediately suspect.<sup>201</sup> This leads to a conception of inquiry as “intelligent experimentation,” where concepts are tested according to their practical consequences for human experience.

Philip Kitcher, for instance, has recently extended this logic into the realm of ethics (an area that was—with the exception of Dewey and James—of relatively little concern to the pragmatists). In this perspective, ethical rules, value systems, and standards of morality are seen as tools that coordinate collective behavior among adaptive people.<sup>202</sup> And much like other tools, they can do the job rather better or worse, depending on the circumstances. However, whether an ethical rule serves the purpose of adequately coordinating depends not on a transcendental truth, but on the practical effect it has on the behavior that is to be coordinated. As such, the truth or otherwise of ethical concepts is determined in ethical practice. In other words, we will be able to test these ideas only *experimentally*, by implementing them and observing (or rather, *experiencing*, since the concept of observation is too narrow) their consequences.

Beyond the classical problems of philosophy, there has also been a similar movement. In the empirical analysis of politics, we have seen a movement towards the incorporation of experimental methodology, both in the academic study of the social sciences and in their concrete application. In the study of complex organizations, this way

---

<sup>201</sup> This aspect of the pragmatist program puts them in close intellectual vicinity of the logical positivists. And indeed the connection between especially the naturalistic strand of pragmatism and logical positivism is closer than people generally assume. See *ibid.*

<sup>202</sup> Kitcher, *The Ethical Project*.



of thinking actually has a long tradition.<sup>203</sup> But only recently have social-scientific experiments become more prevalent.

The use of lab or field experiments has become more prevalent in the study of policy. Development economics especially is an area where the construction of policy interventions as randomly controlled trials has become more prevalent in recent years. Sometimes, this idea gets subsumed under the heading of “evidence-based policy-making.”<sup>204</sup> The motivation behind this approach is that a tightly controlled experimental methodology will provide the most secure information about cause-and-effect relationships within complex social systems.

Beyond the academic study, however, there are actually a number of policy fields where an experimental structure has already been adopted. At this time, it seems that more “technical” fields of policy are more likely to be organized experimentally in this sense. A forerunner here seems to be the European Union.<sup>205</sup> The general methodology is that a central authority sets clear standards, and subordinate units can then try different ways to reach that standard. Subsequently, they will be assessed by their performance and

---

<sup>203</sup> Merton, “The Unanticipated Consequences of Purposive Social Action;” Charles E Lindblom, “The Science of Muddling Through;” *Public Administration Review* (1959): 79–88; Charles E Lindblom, *The Intelligence of Democracy: Decision Making Through Mutual Adjustment*, (New York: Free Press, 1965); Rittel and Webber, “Dilemmas in a General Theory of Planning.”

<sup>204</sup> Abhijit V Banerjee and Esther Duflo, “The Experimental Approach to Development Economics,” 2008; Abhijit V Banerjee and Esther Duflo, *Poor economics : a radical rethinking of the way to fight global poverty*, (New York: PublicAffairs, 2011); Dani Rodrik, “The New Development Economics: We Shall Experiment, but How Shall We Learn?,” *HKS Faculty Research Working Paper Series* (2008); Morton and Williams, “Experimentation in Political Science;” Donald P Green, “On Evidence-Based Political Science,” *Daedalus* 134.3 (2005): 96–100; Nancy Cartwright and Jeremy Hardie, *Evidence-Based Policy: A Practical Guide to Doing It Better* (Oxford: Oxford University Press, 2012); Nancy Cartwright, “What Are Randomised Controlled Trials Good for?,” *Philosophical Studies* 147.1 (2010): 59–70.

<sup>205</sup> Charles F Sabel and Jonathan Zeitlin, “Experimentalist Governance,” in *The Oxford Handbook of Governance*, ed. David Levi-Faur (Oxford University Press, 2011); Charles F Sabel and Jonathan Zeitlin, “Experimentalism in the EU: Common Ground and Persistent Differences,” *Regulation & Governance* 6.3 (2012): 410–426.

the different means are evaluated. The UK Cabinet Office (the body that is central to policy design in the UK system) has recently also restated its commitment to experimental policy-making in an official report, extending the practice to social policy.<sup>206</sup> The US military's shift in macro-strategy in the Iraq War, according to some accounts, may also demonstrate the superiority of experimental tactics in highly complex and uncertain contexts. The change in command over the multi-national forces in Iraq from General Casey to General Petraeus in 2007 was accompanied by a clear break with existing unsuccessful tactics. Before 2007, the war was run in a completely hierarchical way, where data was sent to Central Command and "one-size-fits-all" decisions were then passed down to the units. After 2007 the units (i.e. commanders in the field) were given a larger degree of autonomy in designing their pacification strategies as they wish, and adapt them to local circumstances. The role of central command therefore was transformed to a standard-setter and manager of best practices. The latter strategy, as it turns out, was the more successful one, at least when look at the goals of the US armed forces.<sup>207</sup>

It seems, therefore, that adopting an experimental methodology is an appropriate strategy to find out "what works" under conditions of uncertainty, and across different spheres of inquiry. However, even if experimentation is the appropriate methodology in the natural and the social sciences, and perhaps also with respect to the traditional questions of philosophy, is it justified to implement *policy* experimentally? Even though there is a definite trend towards the use of experimental methodology in the

---

<sup>206</sup> Laura Haynes et al., *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*, 2012.

<sup>207</sup> Tim Harford, *Adapt: Why Success Always Starts with Failure* (London: Little, Brown, 2011).

aforementioned fields, the question is why we should be interested in it in the first place. In the context of this dissertation particularly, we should be interested in whether experimental policy-making can actually be justified according to our normative standard of political authority: pragmatic robust instrumentalism (PRI)?

In order to address this general question, in this chapter I proceed in three steps, with three questions. First, what do we mean by *experimental politics*, and how does it differ from its alternatives? Second, is it even possible to experiment with policy? There are some practical concerns with this; however, there are also deeper questions about whether it even makes sense to think about normative propositions in experimental terms. Finally, even if we can experiment with policy, should we? This is the question of whether it can be ethically justified to create policy (that is potentially coercively enforced) in an experimental way. Recall that this involves the implementation of policy without full knowledge of the consequences (and even possibly in contravention of “best estimates”).

After discussing the nature of experimental politics in section 3, unsurprisingly, I will answer the latter two questions in the affirmative: in section 4, I argue that we *can* experiment with policy—in the sense of the PRI principle we should see policy interventions as resolutions of practical problems, not as verdicts on situation-independent truths. Following this discussion, in section 5 I will defend experimentation against objections from two sides: the position of *optimization*, and the position of *reactionary policymaking*. I conclude that from the perspective of justified political activity, the experimental way is distinctly preferable.

The last section 6 of the chapter leads into the discussion of *democracy* and concrete democratic mechanisms that will follow in the next two chapters, by connecting the outcomes of the analysis in this chapter with the different aspects of democratic decision-making.

### 3. What is Experimental Politics?

Now I have to be a little clearer about what “experimentalism” really means in political contexts. After all, the goals of empirical social science and policy-making can be completely different. Fundamentally, there are two ways we can understand *experimentalism*. On the one hand, we may use experiments to attempt to *verify* certain hypotheses. That is, we attempt to overcome the uncertainty by becoming *more* certain – and the experimental method promises the most secure path to assured knowledge about causal relationships in the social world. This is the traditional understanding of the “scientific method.” On the other, we may want to experiment in order to *adaptively improve*, or as I like to put it, *refine* our practices, beliefs or norms.<sup>208</sup>

In the verificationist sense, an experiment is a designed intervention that should enable to us evaluate the relative plausibility of different hypotheses. Experiments vary from field to field, but they do tend to require some common elements: we need a *hypothesis*, i.e. a conjecture about what the state and the causal structure of the world are; on this hypothesis we ground an *expectation* about the observable *effect* of different

---

<sup>208</sup> The term refinement is taken from Kitcher, and the distinction between the two senses of experimentation is very clearly made by Ansell. Kitcher, *The Ethical Project*; Chris Ansell, “What Is a ‘Democratic Experiment’?,” *Contemporary Pragmatism* 9, no. 2 (2012).

interventions. We may have more than one hypothesis to consider – indeed at a minimum we have to consider the “null hypothesis.” On this basis we design an intervention that aims to produce the desired effect. The reasoning is of course that if the effect occurs, the null hypothesis can be rejected, and vice versa. After that, the intervention must be implemented in a *controlled* manner; meaning the conditions must enable us to distinguish the effect of our intervention from other contingent factors that may have caused the effect. In laboratories that is of course easily done, but outside of those settings, this tends to be done through actively randomized control groups. This model, which is standard in medical research is known as Randomized Controlled Trials (RCTs). For those questions that are of interest for social scientists, possibly “natural” factors could stand in for randomized control if we find ourselves not in a laboratory setting. For example, for accidental reasons a policy intervention may have been targeted at only a random subset of the population, thus “naturally” creating a treatment and control group, and therefore the conditions of an RCT. In contexts with extremely high noise-to-signal ratio – such as the aforementioned complex adaptive systems where multiple causal pathways may exist in parallel, the RCT method may promise reliable hypothesis testing.

However, in the political context getting more and more secure knowledge about different hypotheses is not the primary goal. The primary goal is the gradual adaptive improvement of the rules and norms governing and coordinating citizens’ collective behaviour. Knowing the causal structure of the world may aid in that task, but it is a *secondary* consideration. As it were, if a policy intervention improves on a given set of rules

or norms, we are not primarily interested in *why* it works; what is important is *that* it works.<sup>209</sup>

In the adaptive-refinement conception of experimentalism the transferability and generalizability of the results of the experiment are less emphasized: it cannot be assumed that what works in one place will necessarily work in another. This reflects the complexity of the social world in which policy interventions take place, where multiple factors contribute to a single outcome and contexts vary in a lot of ways.<sup>210</sup> As Chris Ansell mentions, this mode of experimenting is more akin to “design experiments” that are frequently used in educational research, rather than the use of RCTs. It is also probably what Dewey had in mind when he spoke of “intelligent experimentation.”<sup>211</sup> Here, the key is to constantly experiment in a trial-and-error way with the existing policy, and gradually and adaptively refine the functional role of the policy in question. As Dewey put it:

When we say that thinking and beliefs should be experimental not absolutistic, we have then in mind a certain logic of method [which implies] that policies and proposals for social action be treated as working hypotheses, not as programs to be rigidly adhered to and executed.<sup>212</sup>

In any case, on the adaptive-refinement model, *controls* become relatively less important, as the identification of the effect of the intervention is secondary to the adaptive outcome.

---

<sup>209</sup> This logic of experimentalism is developed, among other places, in Kitcher, *The Ethical Project*; Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*; Christopher K Ansell, *Pragmatist Democracy : Evolutionary Learning as Public Philosophy*, (New York: Oxford University Press, 2011); Ansell, “What Is a ‘Democratic Experiment’?.”

<sup>210</sup> On this point see on this point see also Cartwright and Hardie, *Evidence-Based Policy: a Practical Guide to Doing It Better*.

<sup>211</sup> Ansell, “What Is a ‘Democratic Experiment’?.”

<sup>212</sup> John Dewey, *The Public and Its Problems*, (Athens, OH: Ohio University Press, 1927), 202.

On the other hand, what becomes relatively more important here is a working *feedback* mechanism such that the outcome of the experiment has an effect back to the policy. After gathering the information about the observed outcome, i.e. on whether the problem that prompted the intervention has been (provisionally) resolved, there needs to be an appropriate reaction: the experimental intervention is rejected or accepted; and accordingly, our intervention must either be amended or taken back, or can be retained. In scientific contexts the feedback applies to the degree of credence we lend to our hypothesis – but it may be less essential to have this (except in the macro-context). In concrete policy contexts, the feedback however, must lead us to reconsider the policy itself: this is a requirement of adaptation.

#### *Modes of political experimentation*

Now we have to look at how, in a concrete sense, such an adaptive-refinement experimentalism could be realized in the political field. There are several ways we can understand this. First, we can implement one *universal* policy that applies to all citizens, but try to implement it under experimental conditions as far as that is possible. This would involve clarifying the outcome we expect to experience (the “hypothesis”), and of course some feedback mechanism to enable the adaptive improvement. Second, however, we may implement policies *selectively*, such that the new policy applies only to some (ideally randomly selected) subset of the population. This of course makes it much easier to control for different possible influences on outcomes, and thus can give us a clearer indication of whether the intervention was causally responsible for the outcome. For that

reason, this method is also taken to be the gold standard of the RCT mode of experimentalism.<sup>213</sup> Finally, we might just *allow* experiments to take place independently of central control: if we have a federal system, we might allow local and regional authorities to try out different approaches to solve problems, in the spirit of Louis Brandeis' dictum that the American states are the "laboratories of democracy." Finally, we might also allow (and encourage) citizens and civil society organizations to conduct "experiments in living," as J.S. Mill famously calls for in *On Liberty*. In this view, we have to rely on the social dissemination of new norms as the adaptive mechanism. We might call this the *permissive* strategy.

It is clear that the "universal," "selective" and "permissive" strategies to are probably suited to different circumstances. "Universal" strategies are closest to how policy is already implemented anyway, so this mode of policy-making should *prima facie* be relatively uncontroversial. We are already coercively enforcing universal laws, so why not embed them in an experimental framework? Some issues are also not liable to be implemented in a selective way: given the interconnected economy, a macro-economic stimulus or a monetary expansion, for instance, cannot be limited to one geographical location or to one randomized subgroup of the population. On the other hand, it is much harder to disentangle potential causes and effects in a "universal" situation. A "selective" strategy makes this problem of control much easier since one can randomly assign the subjects of the intervention into a treatment and control group. Accordingly, it may be easier to determine the causal effect of the given intervention. However, as mentioned, it

---

<sup>213</sup> Although it is important to note that it does not warrant the inference that it "will work" elsewhere. See Cartwright and Hardie, *Evidence-Based Policy: a Practical Guide to Doing It Better*.



may not be possible to implement policy in this way. It also may bring with it more ethical problems, as discussed below. Finally, the permissive approach also seems relatively uncontroversial, to the extent that it does not force anyone to participate in any experiments; they are essentially voluntary. On the other hand, the less systematic such experiments are, the more difficult it may be to find real solutions.

We have to be clear what we mean by “experimentation” when we discuss its merits and demerits. If we think that a political procedure can in principle be justified if it can be robustly expected to solve political problems, this may call for all three modes of experimental policy implementation, depending on the situation. Which of those strategies works, as it were, is itself subject to experimental inquiry.

#### 4. Can We Experiment With Values?

Now that it is a little clearer what we mean, we can consider a first objection to political experimentation: politics is not concerned only with finding causal connections or the truth about *matters of fact*; frequently, political decisions involve important moral choices as well. It seems obvious how we can understand the idea of experiments with regard to the *factual*, or technological aspects of politics. This is straightforward to understand: if we know the goal we want to reach (like reduce unemployment), but we don’t know how to get there, implementing policy strategies experimentally might be the best way to find out how to get there. And political activity is often concerned with these “technocratic” kinds of problems. Indeed, many examples of experimental policy-making seem to take the desired *goal* of the policy (the “hypothesis”) for granted, and assume the relevant

uncertainty is only about the causal questions of how to realize it.<sup>214</sup> So what this seems to call for is the use of experiments in the *implementation* of policy goals that have been chosen elsewhere. We may call this the “technocratic” sense of policy experimentation.

*Political problems as conflicts in experience*

However, this cannot be sufficient. As discussed in chapters 2 and 3, much of the conflict and disagreement in political activity is about the goals these policies are supposed to achieve. To be sure, there is also a lot of disagreement about the technocratic aspects of politics. But experimentation is supposed to resolve problems of uncertainty associated with differences about normative questions as well.

Understanding political problems, as defined in the Deweyan sense above, as clashes between factual experiences and normative commitments *in our concrete experience*, it seems clear that we should try to change our experiences by affecting the state of our environment in some way. We should bring the social facts into line with our normative commitments: we experience a social phenomenon that strikes us as unjust, so we should try and change the phenomenon. For those kinds of situations “executive experimentalism” may work.

However, it might be more appropriate, in many cases, to adjust our normative commitments instead. Consider the following problem. In most Western societies there are many more same-sex couples openly living together than there used to be, and many

---

<sup>214</sup> Sabel and Zeitlin, “Experimentalist Governance;” Sabel and Zeitlin, “Experimentalism in the EU: Common Ground and Persistent Differences;” Banerjee and Duflo, *Poor economics : a radical rethinking of the way to fight global poverty*.

of them wish to get married. This fact clashes with the normative commitment expressed in standing law that marriage ought to be possible only between men and women—this being based on the normative ideal of the nuclear family. Now, policy resolution of this problem might try and adjust the *fact*, and try and reduce the number of same-sex couples that want to get married: maybe through neural programming or re-education (these are of course, to say the very least, highly problematic). This, however, conceives of the role of political intervention much too narrowly: such a “solution” is only a solution when we look at the problem in isolation from the wider context.

It seems obvious that the more promising solution to this clash would be to adjust the normative commitments expressed in law, and allow same-sex couples to marry as well. One reason is that trying to get same-sex couples to renounce marriage is unlikely to be successful. The more weighty reason of course, is that attempts at re-education or re-programming of this kind violate many *other* normative commitments we have: at the very least, the high value we place on freedom of conscience. A point the pragmatists emphasized is that problem-solving always starts within a concrete social situation, and cannot be conceived in isolation. Inquiry starts in the “middle,” within a set of assumptions that are held constant. In Neurath’s well-worn phrase, we are always “rebuilding the ship at sea.” Problem-solving happens in the middle of our daily lives.

So the political “problem” of the clash between the desires of same-sex couples and the law is not primarily a technical issue. It is a clash between different normative positions: one the explicit assumption in law, and the others that constitute the background normative commitments within which the problem arises.

Therefore this pragmatist account differs somewhat from increasingly popular “technocratic” position (as mentioned above) that experimentation helps us figure out *means*, once we know the *ends*. This also means that this account faces a first crucial problem: Can we even experiment with normative commitments? Can we subject our *ends* to experimental scrutiny as well? Initially, this seems like an absurd notion: how could we even find out whether a policy intervention that resolves a clash of normative commitments “works” or not?

*Value pluralism vs. classical pragmatism*

For instance, assume we hold a value-pluralist position, according to which different normative commitments that people might have are *incommensurable*. That is, there is no common scale (such as general utility) by which we could weigh or trade-off ultimate normative values. Values are not reducible to a common underlying scale. This position is classically associated with Isaiah Berlin and Bernard Williams.<sup>215</sup> But value pluralism remains an active research program.<sup>216</sup> For value pluralists, it is *not* the case that for the different values A and B, A is either better or worse, or of the same value as B. Instead, A and B are *incomparable*.

This would seem to imply that a good “resolution” of the problem of value clashes is not possible. Such resolutions are, in the term of many of the pluralists, “tragic,” in that

---

<sup>215</sup> For instance, see Isaiah Berlin, “John Stuart Mill and the Ends of Life,” in *Liberty*, ed. Henry Hardy and Ian Harris (Oxford: Oxford University Press, 2002).

<sup>216</sup> John Gray, *Two Faces of Liberalism* (New York: The New Press, 2000); Chandran Kukathas, *The Liberal Archipelago* (Oxford: Oxford University Press, 2007); William A Galston, *Liberal Pluralism: The Implications of Value Pluralism for Political Theory and Practice* (Cambridge: Cambridge University Press, 2002).

they involve a necessary loss. Accordingly, we may think that for the political sphere this implies that any normative interventions are prima facie unjustifiable—since there is no standard by which these interventions could possibly be assessed.<sup>217</sup> This implication is a variant of the “No-Standard-Thesis” (NST) that has already been introduced in the context of conflicts of preferences in chapter 3.

However, Robert Talisse has recently argued both against the plausibility of a pluralist conception of value, as well as the view that pluralism implies a version of the NST.<sup>218</sup> For my purposes here I concentrate only on the latter. I shall not get into the question of whether or not pluralism holds as a theory of ultimate normative value. I merely suggest that a commitment to pluralism is consistent with this pragmatic approach to resolving political problems. By implication, if value pluralism passes the test, all other moral stances will as well.

The key argument is that value pluralism as an ethical commitment does not entail that intervention is never justifiable: this would require the further normative principle that strong and unconditional tolerance (or even protection) of all ultimate values *ought* to be the response to pluralism. However, as a pluralist, one presumably has no resources available to assert such a principle: if tolerance is merely one of the incommensurable ultimate values “in the mix,” then its alternatives are equally justified: this includes the view that conflict between pluralist values demands inquiry into how it may be best resolved. The enforcement of the non-intervention principle violates value

---

<sup>217</sup> For that reason, many of the pluralists, such as Gray and Kukathas, advocate a minimally invasive state, i.e. a state that does not interfere with the internal workings of groups and associations, even those that may have illiberal or otherwise problematic values, on the basis that there is no normative ground on which such an intervention could be undertaken.

<sup>218</sup> Talisse, *Pluralism and Liberal Politics*.

pluralism as much as its alternatives, and so the ultimate commitment to pluralism cannot rule out an experimental approach to resolving value conflicts.<sup>219</sup>

However, it does suggest that the experimental approach will be unsuccessful. This itself depends, however, on what we see as “success” in this context. Of course value pluralism means that we will not be able to find the “correct” solution to moral clashes in any transcendental normative sense. So what else could it mean to resolve a political problem?

The key is to understand normative commitments and rules *naturalistically*. Here we can turn to Dewey:

*Moral conceptions and processes grow naturally out of the very conditions of human life. (1) desire belongs to the intrinsic nature of man; we cannot conceive a human being who does not have wants, needs, nor one to whom fulfillment of desire does not afford satisfaction. ... (2) Men live naturally and inevitably in society; in companionship and competition; in relations of cooperation and subordination. These relations are expressed in demands, claims, expectations. (3) Human beings approve and disapprove, sympathize and resent, as naturally and inevitably as they seek for the objects they want, and as they impose claims and respond to them.* <sup>220</sup>

According to this, we see moral commitments as responses to the human predicament of having to live together in societies, and the inevitable clash of “demands, claims, expectations.” As it were, moral theories and moral rules are *motivated* responses to practical problems of social interaction, not results of abstract inquiries into the nature of morality *as such*. Moral theory follows moral practice, in this view, not the other way around.

---

<sup>219</sup> *ibid.*

<sup>220</sup> John Dewey and James Hayden Tufts, *Ethics*, 2nd ed. (New York: Henry Holt and Company, 1932) 343; quoted in Kitcher, *Preludes to Pragmatism* 305.

Given such a Deweyan understanding of “Naturalistic Ethics Without Fallacies”<sup>221</sup> where morality is understood largely as a social technology or tool. As such we can understand the standard to evaluate morality as *functional refinement*.<sup>222</sup> In particular, as already introduced in the last chapter, we can see moral rules (like political rules) as attempts to resolve problems that, following Dewey “grow naturally out of the very conditions of human life.” Moral progress, on this account, consists in resolving problems in more efficient ways.

But what are those problems, and how does this relate to experimentation? Recall the model of political activity as problem-solving mentioned above: resolving clashes between the factual environment as it is experienced by us, and our normative commitments. From this perspective, it seems less problematic to think about how we might assert that a normative policy intervention works or not. It “works” if there is no longer a clash between normative commitments and lived experience.<sup>223</sup>

It is important to note that this does not mean that the normative change would be *optimal* according to some standard external to the problem at hand. There is no guarantee that the normative resolution to a problem would be *fully just* according to some ideal standard, and there is no guarantee that it would even be a final resolution of the problem. There is also no guarantee that a given resolution is the *only* possible resolution to the problem. But then political activity, arguably, is not in the business of delivering verdicts on normative truths; and accordingly the justification of political activity cannot

---

<sup>221</sup> *ibid.* ch. 13.

<sup>222</sup> Kitcher, *The Ethical Project*.

<sup>223</sup> Amanda Roth, “Ethical Progress as Problem-Resolving;” for the classic source, see Dewey and Tufts, *Ethics*; John Dewey, *Theory of Valuation*, vol. 2, 1939.

turn on the extent to which this ideal is reached.<sup>224</sup> Instead, we evaluate political decisions not by their independent truth-status, but by the change they effect with respect to a concrete problem situation. If there is no problem, then there is no cause for a policy intervention, even though from the objective perspective of justice we might still be some way away from the ideal situation (if there is one). If it ain't broke, as it were, don't fix it.

This is, in my opinion, one of the most significant insights of pragmatism, one which is most prominent in Peirce, but has been picked up in significant ways by contemporary pragmatists like Philip Kitcher: the shift from a static to a dynamic model of understanding inquiry and justification. On the pragmatist model, the normative evaluation of inquiry should focus in the justification of a *change of beliefs*, not on the justification of beliefs themselves.

All inquiry is motivated by *doubt*, and not hypothetical doubt, but real, practical doubt from some already existing body of belief. As it were, if there is no problem with our existing beliefs, there is no reason to think about the quality of inquiry. Conversely, if there is a problem, defined as a conflict between normative commitments and the environment, there are automatic standards (the standards of resolution) that come with it. We can say whether a policy has resolved, say, the particular problem of same-sex marriage (a particular clash of normative outlooks and experience of the world), *without answering the question* what the best way to live should be. We can say whether a policy has resolved the problem of unemployment *without* having to answer the question of what the

---

<sup>224</sup> I have argued in chapter 3 for such an understanding of *political activity*.



best way to regulate a capitalist economy is, let alone what a fully just (well-ordered) society might be.<sup>225</sup>

Note also that a commitment to a pragmatist view of politics does not commit us to moral relativism or subjectivism. The question of what a political procedure should do is *prima facie* independent of the question whether its decisions are right or wrong in some more objective sense. One can hold the view that what makes political decisions valuable are objective normative features about them, without the view that political activity should consist in finding as many of these objective “truths” as possible.<sup>226</sup>

If we accept this view of political activity as contextualized problem-solving, we can see that, in principle, we can experiment with normative policies as well. Whether a normative change “works” depends on whether it has resolved the problem that initially motivated it. This is of course not to say that this assessment is *easy*, or even that it is always humanly possible. However, if we think that an experimental strategy is in general the appropriate response to extreme complexity (and I have suggested above that it is), this would make a *prima facie* case for its use in politics.

### *Ends-in-view*

Now, however, we might wonder how exactly this is supposed to work. We have a clash about the final ends of policy: between conceptions of justice, say, or “comprehensive theories of the good” more generally, as Rawls may say. How can the outcome of any

---

<sup>225</sup> Once more, this idea is closely related to Amartya Sen’s work. See for instance Sen, *The Idea of Justice*.

<sup>226</sup> For an interpretation of pragmatic epistemology as a theory of normative epistemology, i.e. a theory about where we should direct our attention, see Kitcher, *Preludes to Pragmatism* 31-33.

experiment sway our opinion about the relative value of final ends? How can we rationally and “intelligently” (as Dewey would put it) reason about the final ends of politics *without assuming some substantive truth of the matter?*

The key is to give up on the clear distinction between *means* and *ends* with respect to political activity. In *Theory of Valuation*, Dewey addresses precisely this question.<sup>227</sup> Dewey’s project here is to extend the logic of the sciences to the study of *value* as well: to come up with a way of reasoning about final ends (“values”) that mirrors the way we reason about causal relationships or matters of fact (what we consider the instrumental means towards those ends).<sup>228</sup>

The first important point is that in our individual reasoning process we do not consider *ends* in isolation of means or problem-contexts, we think about what Dewey calls “ends-in-view,” that is, ends as they feature in action-guiding reasoning. These ends are responses to concrete situations. We may also, in more reflective moods, think about ends in themselves, not in the sense of attempting to realize them, but as an exercise in contemplation. The move from this to having action-relevant ends-in-view, however, involves another step.<sup>229</sup>

---

<sup>227</sup> Incidentally, this work was written as part of Rudolf Carnap’s project of producing a unified encyclopedia of science—the core logical positivist project. Again, we can see the close relation between pragmatism and logical positivism.

<sup>228</sup> See Dewey, John. *Theory of Valuation*. International Encyclopedia of Unified Science, Vol. 2, (Chicago, IL: Chicago University Press, 1939); see also the insightful discussions in Roth, “Ethical Progress as Problem-Resolving;” Henry S Richardson, “Truth and Ends in Dewey’s Pragmatism,” *Canadian Journal of Philosophy. Supplementary Volume 24* (1998): 109–147.

<sup>229</sup> In the same way that we may occasionally think about how it would feel like to fly unaided, or to be king of the world. These are *final ends*, but they do not usually feature in the determination of our actions. *ibid.*, 117.

Second, it is not true that we first determine which goal we have and then, once that is fixed, determine how to get there. Rather, the means necessary to a goal decisively influence our evaluation of that goal. The intuitive point here is that the means necessary determine whether the end is worth having in the first place. In this context, Dewey himself uses a rather fanciful example involving the story that roast pork was first discovered when a stable with pigs inside accidentally burnt down. Now, if the only method to get roast pork was to burn down a wooden stable each time, this should decisively influence our desire (“end”) for pork: this would obviously not be a goal worth having.<sup>230</sup>

But we may come up with a more political example: Imagine the only way to get full equality of resources would be to “level down”—that is, to make everyone worse off in terms of their property holdings in order to create equality, albeit at a lower overall level.<sup>231</sup> This fact may lead us to *question* our chosen end of full equality of resources: we may conclude it is not an appropriate end-in-view.

Indeed, what seem to be means themselves may take on the role of “ends-in-view” in our thought process: on the path to a final outcome we have to execute certain intermediate steps, and to the extent that they are action-guiding they become “temporary” ends-in-view. And as such they may of course conflict with final ends I may have: they need to be evaluated against those other ends. Indeed with respect to many activities—especially also social activities, we are *primarily* thinking about ends in view and evaluating their respective merits. For example, consider chess: the ultimate end is of

---

<sup>230</sup> Roth, “Ethical Progress as Problem-Resolving” 392.

<sup>231</sup> See for instance Nils Holtug, “Egalitarianism and the Levelling Down Objection,” *Analysis* 58.2 (1998): 166–174.

course winning the game, but I cannot go about a chess game trying to achieve this directly (chess is too *difficult* for that): I have to set myself heuristic ends-in-view, such as “gain control of the center of the board.”<sup>232</sup> My individual actions are evaluated in light of these ends-in-view, simply because I cannot evaluate individual moves according to the standard, “win the game.” And of course, the different ends-in-view may conflict: in a specific action-context “gain control of the center of the board” may come in conflict with “trade pieces only against pieces of lower relative value,” or indeed with the ultimate end “win the game” (depending on the opponent’s behavior).

This may seem like a roundabout way of stating the intuitive point that sometimes, the actions necessary to realize one of our goals (which results from normative commitments, for example) may itself be conflicting with another normative commitment we have. Equality may not be worth leveling down, social order and peace may not be worth cracking down on freedom of expression, and so on. However, more importantly, what this shows is that there really is *no sharp distinction* between means and ends when it comes to the normative evaluation in practical situations—the context with which we are concerned here. With respect to specific problem-contexts ends can and have to be evaluated in the same way as “means.”<sup>233</sup>

Now we can return to the question at hand. Those concrete problematic situations which call for a resolution—conflicts between normative ideas actual experience—can be represented as conflicts between different ends-in-view. As such, of course, this conflict is nothing out of the ordinary and in fact to be expected. Since we always as a matter of

---

<sup>232</sup> I owe this insightful example to Jonathan Bendor.

<sup>233</sup> On the same issue of the indistinguishability of factual and normative claims, from another angle, see also Putnam, *The Collapse of the Fact/Value Dichotomy: and Other Essays*.

course attempt to resolve these conflicts in our daily activity, this should also not seem insurmountable. And clearly, if we see concrete problems not as conflicts between isolated commitments of intrinsic value, any information gained through experience and experiment can be valuable for us to evaluate and resolve that conflict.<sup>234</sup> If a clash between ends-in-view is seen as a problem of choosing under uncertainty—much like actual political disagreement—then the experimental method may indeed be the proper response.

## 5. Should We Experiment With Values?

Now that I have given some reasons to believe that experimentation in politics is *possible*, we can move on to the question of whether experimenting with policy can ever be desirable or justified. In this section I consider a few objections. My answers each time will follow roughly the same pattern: if we think policy intervention is *ever* justified, it is justified to implement it experimentally as well.

Prima facie, this seems an odd conclusion: after all, when we experiment with policy, we seem to imply that we are not sure of the benefits of the policy. How, at the same time, can we claim authoritative status for the policy?

The main salient objections to an experimental approach to politics can be usefully subsumed under three headings; which for this section I will borrow from Albert

---

<sup>234</sup> Compare also MacGilvray, “Experience as Experiment: Some Consequences of Pragmatism for Political Theory.”

O. Hirschman's *Rhetoric of Reaction*: (a) perversity, (b) futility, and (c) jeopardy.<sup>235</sup>

Hirschman focused on conservative arguments against progressive policy agendas more generally. I will not address those arguments here. Instead, I will focus on three main objections to *experimentalism*, which do not necessarily match up exactly with the arguments in Hirschman's book.

Throughout, I will contrast the experimental approach to policy-making with two alternative modes of policy-making: *optimization* and *reaction*. These are of course ideal-types, much like the experimental strategy itself. However, in the political context and basic normative framework I am using, ethical justifiability is a *relative* concept—recall the discussion of how to establish justifiability in chapter 3. Any justified exercise of authority, in this model, is only justified relative to its alternatives.

Optimization we have already encountered above: this refers to policy-making on the predict-and-act model. This involves estimating the probabilities of unknown parameters to the best of one's knowledge, taking those estimated probabilities as objective and then maximizing the expected payoff through the policy choice. This strategy has also been called “eggs in one basket,”<sup>236</sup> since it essentially represents a bet of the “whole house” on the future state of affairs being in a certain way.

The alternative view is the polar opposite: *reactionary* politics, denies that we should ever do anything *unless we are absolutely sure of the consequences*, because the slightest risk outweighs the potential benefits of departing from the status quo. This is of course an extremely risk-averse strategy of policy-making, but it is not in principle irrational. I term

---

<sup>235</sup> Albert O Hirschman, *The Rhetoric of Reaction* (Cambridge, MA: Harvard University Press, 1991).

<sup>236</sup> I am indebted to Tom Slee's blog for this term: <http://whimsley.typepad.com/>

this strategy *reaction*, rather than conservatism; these notions have to be clearly distinguished. Conservatism seems to me characterized by a desire to *conserve* certain specific values and practices (especially those that are traditionally received), these practices are then accorded special (maybe even lexically prior) value just because they are in place, even though alternatives may function better. The conservative values these things worthy of preservation because people have formed particular attachments to *them* even though there is no independent rational reason why they are to be preferred.<sup>237</sup> Therefore, conservatism *may* imply a reactionary political stance, but not necessarily so. As it were, if circumstances change decisively, conservatism may require *decisive and proactive action* to ensure the goal of conserving whichever values there are. In the oft-quoted phrase from *Alice in Wonderland*: “It takes all the running you can do, to stay in the same place.” Conversely, one may adopt a reactionary position without a fundamental commitment to conservatism: instead, for instance, because of extreme risk-aversion.

Recall once more the normative framework of pragmatic robust instrumentalism: this involved the claim that it is in every reasonable person’s (avowable) interest that an effective and robust political problem-solving mechanism be in place. This was then interpreted as requiring an *experimental* method of problem-solving. So what if we can show that being subject to experimental policy implementation is actually *not* in everyone’s reasonable and robust interest?

---

<sup>237</sup> this view largely follows Ted Honderich, *Conservatism* (London: Pluto Press, 2005); see also Corey Robin, *The Reactionary Mind: Conservatism from Edmund Burke to Sarah Palin* (Oxford: Oxford University Press, 2011).

### *Perversity*

The “perversity” objection against progressive policies in Hirschman’s original version consists in the claim that the policy will achieve the *reverse* of the intended objective: raising the minimum wage, for instance, will make low-paid workers *worse off*, since they are liable to lose their jobs. Here I consider a similar kind of objection: that trying to improve the concrete experience of humans through experimental policy-making will actually *harm* their interests.

There are two ways to interpret this claim: first, the experiment may of course *go wrong* and harm people’s interests in that way. This will be discussed below under the heading of *jeopardy*. Second, however, the practice of experimenting *itself*—i.e. regardless of its consequences may harm people under it. The practice of problem-solving, as it were, exacerbates the problem it is supposed to solve.

Now the first thing to note is that the perversity objection applies to *optimization* approaches as well: if it is problematic to harm people by exposing them to new policies that are implemented experimentally, it would also be problematic to expose them to ones that are implemented in an optimization way. Therefore, we have to look at whether the perversity objection implies a commitment to reactionary policy-making: that experimental problem-solving of any kind is actually not a legitimate use of state power.

This position may be most closely associated with Chandran Kukathas’ work on value pluralism and its implications for the limits of legitimate state activity.<sup>238</sup> The issue of value pluralism has already been discussed above, of course, and I will not here go into

---

<sup>238</sup> Kukathas, *The Liberal Archipelago*.



a discussion of all the arguments in the value pluralist arsenal. Beyond the basic meta-ethical assumption of pluralism, one of Kukathas' key arguments is that people may be *harmed* by being exposed to alternative ways of living from their accustomed one. If this is true, then of course any attempt to *improve* a problematic situation through “experiments in living” or more mundanely through social policy is possibly unjustified: how can I legitimately impose the harm of experimentation on non-consenting citizens.<sup>239</sup>

Now we have to understand how this harm is supposed to come about. I have already discussed above that pluralism and the thesis of the incommensurability of values do not necessarily imply the normative principle that all value commitments *ought to be* tolerated: if different value commitments cannot be compared to each other, there is no sense in which a change from one to the other constitutes a *harm*. There is no clear sense in which anyone is made worse off. A frequently cited example from pluralist literature states that Shakespeare and Mozart are incommensurable. Accordingly, one may wonder how anyone is made worse off by being made to choose Mozart over Shakespeare, or the other way around.

So, the perversity objection requires the further assumption that coercively exposing people to different ways of living—different *from the one they happen to have espoused*—harms them. Kukathas cites two ways in which such an argument may be made: first, some people may be *unable* to reconsider their fundamental normative commitments; there is no external viewpoint available to them. Examples may be the Samurai, who have deeply internalized a code of honor, or devout Muslims who are unable even to

---

<sup>239</sup> Similar points are also made in John Gray, *Mill on Liberty: A Defence*, Second Edition. (London: Routledge, 1996).

reconsider the basic tenets of Islam.<sup>240</sup> These people follow their *conscience*, and this faculty, according to Kukathas, is impenetrable to rational considerations. Thus, these people *have no interest in being able to revise their ends*. The second argument centers on the idea of people who are *happy* with the unquestioned life they live. A reason may be, for example, that they will feel alienated from their communally held lifestyle if they are being made aware of other ways of living. These people, in contrast to the Samurai, *have an interest in not being able to revise their ends*.

I have already discussed above that pluralism and the thesis of the incommensurability of values do not necessarily imply the normative principle that all value commitments *are inviolable*. If different value commitments cannot be compared to each other, there is no sense in which a change from one to the other constitutes a *harm*.<sup>241</sup>

So, the perversity objection requires the further premise that coercively exposing people to different ways of living – different *from the one they happen to have espoused* – harms them. Without going too deep into this, we may for now accept that people with very basic identity-defining religious or moral beliefs (impenetrable to rational calculations) may receive some degree of harm when they are subjected to a new experimental policy.<sup>242</sup> This is not a defect, as “the unexamined life may well be worth living.”<sup>243</sup> Let us grant this.

---

<sup>240</sup> Kukathas, *The Liberal Archipelago* 47-48.

<sup>241</sup> Talisse, *Pluralism and Liberal Politics*.

<sup>242</sup> Kukathas, *The Liberal Archipelago*, 47-8.

<sup>243</sup> *ibid.*, 59.

Now, however, this implies that only the reactionary mode of policy-making is justified *only if* for these people no conflicts of ends-in-view are possibly going to arise. This is of course very unlikely, unless one of two assumptions hold: either their most fundamental commitments and their practical lives are strictly compartmentalized, such that what goes on in one does not affect the other; or alternatively, that the set of their fundamental commitments is essentially *complete*, that is, it will give definite guidance for all possible situations they might encounter.

The former assumption requires that the actions one might have to perform in social life (the “ends-in-view” one may have to endorse) never impinge on the fundamental normative commitments (the “final ends” one has), while the latter assumption requires that there should be a definite end-in-view for every situation recommended by the ultimate commitments one has. I just want to suggest that neither of those assumptions are realistic, unless these people happen to live in a static and autarkic society where everyone already shares those fundamental commitments. But this is clearly a special case. If conflicts of ends-in-view may arise, they need to be resolved in some way – and why would the experimental method of resolving them be particularly problematic (especially if a dogmatic way of resolving them may not be available)?<sup>244</sup>

In the social sphere this problem is of course only exacerbated: conflicts in concrete experience may arise much more frequently between people of different normative commitments than between one person and their environment. Therefore, the perversity objection as I have stated it does not immediately imply a rejection of

---

<sup>244</sup> This reflects Peirce’s “method of authority” of fixing beliefs, which, as he points out, will work for a while, but only until it runs into practical conflicts. Peirce, “The Fixation of Belief.”

experimentalism: even people who have a fundamentally conservative attitude to ultimate moral commitments may have to endorse that there should be a political mechanism to resolve these practical problems robustly.

### *Futility*

In Hirschman's original sense, the "futility" objection refers to the view that people/the economy/values/the human condition/etc. are essentially *stable* across time, or that they change independently of human action. Examples of theories that ground this objection may be "iron laws" of various kinds, especially iron laws of history or "national character."<sup>245</sup>

All political intervention does is change surface conditions, but policy interventions are unlikely to change anything meaningful, neither in a progressive nor in a regressive direction. Consequently, experimentation is unnecessary, and actually a waste of resources. Presumably, the futility view implies that the task of politics ought to be contained to the mitigation of the issues that arise from these basic facts.

Applied to the experimental context, we might say the futility objection focuses on the opportunity cost of all those experiments that do not work out. While there may be some benefits to experimenting, this loss may be so high that it cancels out the value of all those benefits. Especially from the *optimization* perspective, this becomes a problem. Recall that we expect the experimental method to be relatively *inefficient* in the progress it makes. Experiments may occasionally (or frequently, for that matter) lead down blind alleys. But

---

<sup>245</sup> Hirschman, *The Rhetoric of Reaction*.

beyond that, the Deweyan perspective outlined above suggests that suboptimal solutions are fine – there is no push toward maximization. Progress is always *from* somewhere, not *towards* an ideal. This again implies that an experimental political system may get stuck in a suboptimal situation, while expending lots of resources on failed experiments. How can we justify this cost?

I just want to note here this opportunity cost only arises if there really is a foregone option: if optimization is a real possibility. This again depends on some degree of certainty both about which are the right ultimate values, how they are to be traded off, and how we can bring them about with the policy instruments available to us. Now I do not want to exclude in principle the possibility that we can gain this certainty. It may be that we will some day find the definitive theory of justice, for example. However, in the context of public justification, it is of course difficult to premise the justification of a particular optimizing policy on the assumption of some substantive “truth” of the matter. If we allow some degree of disagreement and uncertainty about *ends-in-view*, and how they can be traded off with one another, the optimization strategy, while desirable in theory, is not justifiable. Hence, it is not available as an alternative when evaluating an experimental mode of policy-making.

In short, while experimentalism is specifically predicated on the assumption of uncertainty, optimization strategies require the opposite assumption. Nevertheless, there still is an important aspect to the futility objection: experimentation can of course not go on forever: at some point one has to settle on a practice that experimentation has determined, in order to *exploit* the gains from experimentation. If I spend all my time trying to find out which is the best car to buy I will forgo the gains from actually using it.

At some point, the loss incurred from not completing the purchase will be higher than the additional benefit I may get from *definitely* having the best car. Enthusiastic experimenters should keep this in mind.

However, it is important to note that experimentation here is not actually in the service of optimization: on the pragmatic problem-solving perspective, the experiment can be considered a success when “the supposed good solves the problem which prompted our inquiry in the first place.”<sup>246</sup> The experimentation-exploitation tradeoff is therefore less acute: we have a relatively clear stopping rule – as opposed to experimenting in order to find the *optimal* way to do things.

### *Jeopardy*

Now let me consider the last potential objection, which seems to me also the most frequently cited one in this context. The *jeopardy* objection is close to the *perversity* objection, but not exactly the same. Here it is: policy-making under uncertainty typically involves the imposition of *risks* on individuals. Therefore, if it involves risks, experimenting with people (and by analogy also with policy, which has direct effects on people’s lives) is morally objectionable, especially if it is done coercively and without explicit consent. Even if it is on the road to progress, typically, there will be some who will “lose” from any given policy. Therefore, there is the question: can experimenting ever be justified to the *losers*? For instance, consider the following hypothetical example: a government decided that because of its high unemployment its industry needs to be modernized, to create more

---

<sup>246</sup> Roth, “Ethical Progress as Problem-Resolving,” 391.

wealth and more secure jobs. The government decided on an experimental strategy and invests heavily in high-tech. This is successful and new thriving sector is created. However, this replaces the old types of manufacturing, and the new sector does not require unskilled workers anymore. Consequently, many of them lose their jobs. Thus: can the modernization of industry be justified even to *those workers who lose their job?*

This argument has some intuitive appeal. Consider what seems to be an analogous case: drug trials. The use of randomized controlled trials in testing new drugs, for instance, is absolutely indispensable. We do not and should not accept any medicine that has not been tested on humans. The effects of new chemicals on the body are too incompletely understood to allow us to give any *ex ante* confidence to hypotheses of the effect of drugs. This means that there is an enormous epistemic benefit to adopting an experimental form of inquiry when it comes to allowing or restricting the availability of drugs. Yet, medical trials involve grave risks to the test subjects, including a risk of serious incapacitation or potentially even death. However, the key aspect of drug trials is that they are *voluntary*: participants have to give explicit consent. We are, quite rightly, horrified by forced medical experimentation.<sup>247</sup> But politics is by its nature *coercive*, and people are under certain political authorities—and have to obey or are made to obey laws—generally regardless of whether they consent to that particular law or not. Therefore, if we implement experimental policies, we basically force citizens to comply with a new policy which is incompletely understood. Can that be justified? The first thing to note here is of course that this objection only applies to the *universal* strategy of

---

<sup>247</sup> At least in reasonably ideal circumstances: poverty, a lack of otherwise available healthcare, or being a soldier, may effectively *force* people to undergo medical trials because they lack other options. In societies with these characteristics, we may already consider that people are as a matter of course undergoing medical experiments involuntarily.

implementing experiments. Only if the experiment is actually implemented through *coercive* policy that it is impossible to opt out of. *Selective* experimental strategies of implementing policy could possibly (with some exceptions) be designed on the basis of voluntary participation—for example, we could ask people whether they want to participate in a new scheme to prepare them for the labor market. With experimental local school projects in the United States this strategy is sometimes used: and the fact that those experimental projects are often hopelessly oversubscribed suggests that people are in fact not reticent to participate.<sup>248</sup> Of course, this approach may not always work for all policy fields. Finally, on the least invasive model of experimental politics, allowing or encouraging people to experiment with new forms of living, such as is John Stuart Mill’s ideal, we are not coercing anyone to do anything—indeed, quite the opposite.

However, the fundamental assumptions of the jeopardy objection may hold for large areas of policy; and the medical experiment analogy may also be an apt one. Politics is, in a very real sense, concerned with and impacts on people’s lives. My claim is, however, that even coercive experimental policy-making, whether universal or selective, can in principle be justified.<sup>249</sup> The answer lies in once again considering what the alternatives to experimental policy-making are: “optimization” acknowledges that coercive policy interventions may in principle be justified, but argues that one may not *experiment* with policy. One should just implement universal laws according to the best available judgment (moral and/or factual), and then keep them indefinitely.

---

<sup>248</sup> This depends, of course, on the quality of the status quo.

<sup>249</sup> This is not to say that *all* experiments are necessarily justified, of course. We may still have standards about legitimate state coercion—but this of course holds across all forms of policy-making.



Now, if we believe optimization is in principle acceptable, my claim is that then experimental politics is also acceptable. If we believe the former is acceptable, we also have to accept the latter. Similarly, if it is warranted to implement a policy that is coercively enforced on the best available evidence, then it is equally warranted to implement it with experimental controls and under experimental conditions. As it were, we get all the benefits from regular methods of policy-making while being prepared for the worst. The implication of this condition, of course, is that experimental policy-making should not be done *lightly* either: experiments ought still to proceed on the best available evidence and for good reasons only, although the standards may of course be lower for the less intrusive forms of experimenting.

Now let us consider the objection from the reactionary side: that the danger of imposing risks on people through experimentation means that it is not justified. If we take this path, as I argue, we are probably illicitly privileging the status quo. We are likely to be suffering from *status quo bias*. Just because we have something now does not mean that nothing better is available, and more crucially, it does not mean that what we have now is morally acceptable at all. This is an important point. Note that reactionary politics does *not* mean we are not coercing anyone at the moment, it just means that we are coercing them to conform to the *current* system of laws rather than any other future one. And, importantly, the current system of laws and policy may be just as bad, or even worse, than the outcome of our experiment.

However, we can test whether we (or our political institutions) are suffering from this status quo bias quite easily through the ingenious device of what Bostrom and Ord

call the *Reversal Test*.<sup>250</sup> If we find ourselves in a situation where we are considering the imposition of risk of harm on some individuals, we can perform this simple thought experiment: imagine we are in a state that is much improved from the current situation. Would we actively want to return to our current state? If we answer no, we suffer from status quo bias. To take Bostrom and Ord's example: Imagine it is possible to improve people's intelligence through performance-enhancing drugs. Now let's say, as surely many do, that we think that this is morally perverse, and that we should keep intelligence levels as they are now. Now they ask you to imagine the following situation: a chemical accidentally leaks into the water supply, raising everyone's I.Q. by 10 points. Would we then support actively inflicting brain damage on people to reduce their I.Q. back to present levels? To the extent that most people say no, therefore, this shows that our initial judgment was informed by status quo bias. There is no reason to suppose that our current predicament is the best unless we say yes to the reversal test.

What does this mean for the political context? It means that a selective roll-out of policies is unjustified *only* if generally, we think it is problematic that policies affect only subsets of people in general. For instance, as a matter of fact, frequently policies decided at the country- or federal level are implemented in a staggered way across municipalities or states. Some cities or counties tend to be a little earlier than others in implementing policy. Sometimes, states can even opt out of federal policies. If we do *not* think that such a situation is problematic, we should not think that a randomized controlled roll-out of

---

<sup>250</sup> Nick Bostrom and Toby Ord, "The Reversal Test: Eliminating Status Quo Bias in Applied Ethics," *Ethics* 116.4 (2006): 656-679.

policy is problematic. As it were, the fact that it is done voluntarily and consciously does not change the situation in a structural way.

However, this may not be sufficient: the people who *are* actually harmed may still have a reason to complain even though from the perspective of society it may be justified to experiment. Consider, for instance, the factory workers in the example above, who lose their jobs as the country moves towards high-tech manufacturing.

Again, we can consider the alternatives to experimental policy-making here. If we believe *optimization* is justified, we also should believe experimentation is justified as well. Implementing the modernization policy straight-out and universally, will create just as many losers as implementing it experimentally. Indeed, the latter may be much preferable: for instance, the high-tech initiative may, for testing reasons, be initially confined to only one limited area (think of a special economic zone like Shenzhen in China).

The other alternative to these two modes was the reactionary one: do nothing in order to minimize the risks from change. Again, this is justifiable only if the status quo is better than the expected outcome of the intervention. This means that the “insider” unskilled workers keep their jobs, but otherwise the unemployment remains high. Remaining in the status quo therefore needs to be justified to those potential workers as well. Thus, reactionary politics is acceptable only if we would vote yes on the reversal test scenario: imagine industry modernizes without any active government intervention. Would we be justified in actively destroying the new high-tech sector in order to protect the old manufacturing jobs?

Therefore, it is not the case that doing nothing is not imposing risks or likely harm on anyone. Indeed, if minimizing risk to individuals is our concern, we—once again—may well need some *extensive interventions*. Unless we pass the reversal test, there is no reason to assume that the current level of risk is necessarily acceptably low. There is, in the technical sense I have been using the term, a *political problem*: a clash of interests (this time of insiders and outsiders) that requires resolution on the social level. This shows that *reaction* is not automatically the appropriate response to such a situation. Instead, further inquiry should be used to weigh the different risk impositions on different sectors of society in a way that may resolve the issue.

This distinction however maps on to the fundamental difference between Roosevelt and Adenauer that is alluded to in the very beginning of this chapter. The difference is of course that during Roosevelt's presidency, remaining in the status quo was very much risk-laden and imposed heavy burdens on most people. In the midst of the Great Depression, any risk associated with potential economic improvement seems minimal in comparison to the status quo. From the perspective of the German *Wirtschaftswunderjahre* (years of the economic miracle) that fell partially into Adenauer's tenure, on the other hand, the status quo looked pretty good, and the downside from possible failed experiments seemed acute. Now, of course, one should not overstate the risklessness of the Adenauer years: especially one may question whether an apparent stability does not hide further systemic risks, and therefore warrants some experimentation (albeit limited).

Another version of the jeopardy argument would be to argue that sometimes, the risk of any policy intervention is so great that it ought not to be done in principle. Given

the complexity and potential criticality of adaptive systems, we might think, any policy intervention may, with some non-negligible probability trigger a catastrophic consequence. I fully accept this observation; however, an *experimental* approach to policy-making (to the extent possible) is actually the *appropriate response* to high systemic risk: as opposed to, again, *optimization* or *reactionary politics*.

Consider this simplified example: air traffic control is an area with small, but extremely high-stakes risks. The likelihood that there will be an accident is low, but when there is one, the loss of life is tragic, and should absolutely be prevented. Assume (realistically) that we do not know which system of guaranteeing the safety of air travel works best. We can implement the system that according to our best estimate will work. But we might not be right—we would need to implement experimental conditions to find out. But how could we possibly experiment with air traffic security? We cannot implement new procedures and count how many planes crash. This is of course unacceptable. The regulation of nuclear power plants and defenses against natural disasters have the same problem structure. It seems that we cannot really experiment here, we have to stick with what we have got, whether it is adequate or not.

However, as it turns out, we *can* (and do) employ experimental methods even in those high-stakes areas. The key lies in periodically subjecting the safety procedures to testing and counting *near misses* as *failures*.<sup>251</sup> Even if we cannot count airline disasters in order to weigh different possible ways to organize air traffic control, we can count the number of times a crash was avoided only by, let's say, only a single “last-line” safeguard

---

<sup>251</sup> Charles F Sabel and William H Simon, “Minimalism and Experimentalism in the Administrative State,” *Georgetown Law Journal* 100 (2011): 53–93.

measure. This is in fact how this system works, as Sabel and Simon suggest. In high-stakes areas, thus we may conduct these *quasi-experiments* to find out reliable information.

In political contexts, the near-miss approach to evaluating experiments may be appropriate with respect to many areas: consider financial regulation, terror prevention, or environmental protection. It may not be appropriate for all areas, but we can see that the appropriate response to low-risk, high-stakes situations is *not* to refrain from any attempt at improvement.

Thus, I hope to have shown that an experimental approach to policy-making can be defended against most of its key ethical objections. Given the need for an experimental approach to politics outlined in the first section above, this alone seems a somewhat reassuring conclusion. In the next section I therefore move on to consider a different angle of this whole problem. Given the definition of experimental politics I have offered, does it have any real implications on how we should evaluate democratic institutions? Is there anything in particular to which an experimental approach to policy-making actually commits us?

## 6. Conclusion: What Does This Spell for Democracy?

The foregoing discussion may have led the reader to believe that experimental politics has very unclear contours. It seems all things to every man. After all, if it is such a common-sense view, and if it is compatible with almost all intuitions about policy-making except the most extreme ones, it does not seem much of a substantive theory. One may also

wonder whether my characterization of the alternatives, *optimization* and *reactionary politics*, has really been fair. Partially this has to do with the fact that in my view, experimentation just *is* common sense; it has few apparent downsides compared to other policy-making methods. The fact that it is not seen as the default mode of implementing policy may stem from a period where the data-gathering necessary for experimental politics was impossible for technological reasons.

Furthermore, however, this experimental mode of policy-making seems (at least *prima facie*) compatible with all kinds of forms of policy-making. This is of course a problem if I want to argue that an experimental mode of policy-making can form the basis of an argument for justified democratic authority. In particular, this is problematic because it seems we do not yet have a critical perspective—we do not yet know *which* mechanisms of democracy can be justified on an experimental basis. Which of them can be *robustly* expected to deliver the required functions of an experimental form of policy-making?

The basic outline of the argument has already been offered in chapter 3 above, but we do not yet know what precise mechanisms will actually do so. The next two chapters will spell out the precise answer to this. Nevertheless, let me briefly address one issue here. One might well think that experimental politics as I have just laid it out smacks of an elitist view of politics, with a group of purported experts sitting on the top of the political hierarchy, trying this or that policy on people without consulting them much. It is true that there might be an element to this: it is not *in principle* objectionable to an experimental theory that experts should *design* policy interventions. Indeed, for an experimental model it does not really much matter *who* is making the policy at all, as long

as there is appropriate feedback available to enable functioning experimentation. It is important to note, however, that the role of the so-called “expert” is very different in this model: an expert is not someone who is likely to get it right: as per the assumptions of the experimental model, political problems are so complex that it is unlikely that experts will agree, let alone that we will be able *ex ante* to identify which one of them will get it right. I will refer to the deep and widespread disagreement among economics “experts” regarding the current financial crisis if anyone should doubt this argument. In the experimental model, expertise is instead defined as awareness of the limits of political knowledge and the right *methodological* skills to implement experimental policy.

In addition, I would remark that that is *already* how the actual laws are written and designed and we do not consider that much of a problem. The general citizens usually has little insight into the actual process of composing legislative proposals, and while debates in the chamber may be public, the question which staffer has actually written which provision is difficult to answer. Indeed, the requirements of the experimental model to *specify* and *clarify* the expected (measurable) consequences of a policy intervention would arguably *improve* the actual control people can exercise over the legislative design process, in that it would bring this process out of obscurity to some extent.

Beyond that, as already pointed out above, an experimental system of policy-making will only work *as part of a democratic system of politics*. Since experimental policy-making depends on input as to which problems to resolve, and whether or not they have been resolved, there must be *meaningful* ways to articulate these things. The focus of democracy, however, will shift somewhat: from the frequency or directness of popular involvement towards the *clarity* of the people’s judgment and communication. If the direct



influence of the people on to the *policy-making* process becomes less important, the *policy-evaluation* influence of the people must become much stronger. Last but not least, an experimental system of politics depends on the experimenters' goodwill; and there is no good reason to assume that they will necessarily keep honest unless subject to electoral control. Hence, the experimental politics model can give us clear guidance on the difficult relationship of democracy and so-called expertise.

As it stands, therefore, this chapter gives a modest defence of experimental strategies of policy-making against some moral and factual objections. Especially if we conceive of politics in a Deweyan mold as problem-solving through progressive adjustment and reconciliation of concrete experiences of people and the commitments of normative systems, the idea of experimentation becomes more plausible. This view of politics, it must be said, sees political activity as essentially as a different kind of beast than moral philosophy or similar intellectual pursuits. Similarly, political activity should be evaluated according to its own standards, and not according to the standards of moral philosophy writ large.

This concludes the theoretical part of the dissertation: A long way of establishing the first premises of what I have called the *pragmatic argument for experimental democracy*, and which is reproduced here:

- (1) Only those political procedures are justified that can reasonably be expected to be robustly better than their alternatives at progressively overcoming political problems as they arise.

- (2) Given the extreme uncertainty that surrounds political problems—in the sense of both disagreement and ignorance about the problem structure and/or possible solutions—*experimental* methods of policy-making can reasonably be expected to progressively overcome political problems, better than alternative systems.
- (3) Experimental methods of policy-making have certain functional requirements.
- (4) (Some) democratic political structures (those I will call elements of “experimental democracy”) are best suited to fulfill the functional requirements of an experimental model of policy-making.
- (5) Institutions that fulfill the functions of experimental democracy enjoy legitimate political authority.

Institutions that fulfill the functions of experimental democracy enjoy legitimate political authority. The next general part of the dissertation opens the box labeled “democracy,” and looks more concretely at which institutional mechanisms of democracy actually can be expected to function according to this ideal. Thus, the next two chapters will complete the argument for experimental democracy, and show how we can use this argument to evaluate our own institutions, as well as their alternatives.

## Part II

---

### A Theory of Experimental Democracy

---

## Models of Democracy

### 1. Introduction

The previous part of the dissertation has discussed and endorsed a pragmatic, experimental conception of *political* legitimacy. This second part of the dissertation addresses the question what an experimental form of *democracy* in particular might look like, and indeed why and in what sense this conception of legitimate political authority calls for democracy at all. That is, what general structure would a democracy have to have, that satisfies the functional requirements of the experimental model of politics? Therefore in this chapter and the next, the pragmatic argument for democratic authority will be completed.

If you recall, the stated goal of the dissertation was to come up with a normative theory that gives us a critical standpoint from which we can differentiate different institutional embodiments of democracy. I look at this question at a more general level in this chapter, and in greater detail in the next one. The present chapter is concerned with the question which general conception or “model” of democracy most closely accords with the experimental ideal, while the next one will delve deeper into the concrete

institutional arrangements that would make an experimental form of democracy possible. The basic argument of this chapter is two-part. first, there is the negative point that no particular first-order decision mechanism of democracy can robustly and reasonably be expected to regularly deliver the problem-solving results we are looking for. As I show in this chapter, we simply have no reason to believe that those mechanisms will robustly deliver the problem-solving outcomes we want. Rather, successful problem-solving may require institutional experimentation with first-order decision procedures as well.<sup>252</sup> Some problems may be adequately addressed by the plenum, some by referendum, some by a small deliberative committee—and yet other may be better addressed by an independent body, of experts, bureaucrats, or judges (or central bankers). In more general terms, we might say that decisions are best made by situation-specific *decision networks* that may involve representatives, stakeholders and independent expertise.<sup>253</sup>

*Second*, however, there is the following positive point: I argue that these experimental institutions have to be embedded *within* a democratic structure, in order to make experimental decision-making possible, and to ensure that the political system *robustly* responds to problems and is forced to address them. Remember that according to the PRI principle a political system can be justified to reasonable citizens only if they have

---

<sup>252</sup> A very closely related argument is made in Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*. The authors pursue this project even more explicitly in forthcoming work, see Jack Knight and James Johnson, “Democratic Experimentalism,” Midwestern Political Science Association Annual Meeting, 2013, 1–40.

<sup>253</sup> Chalmers, *Reforming Democracies*.

a reason to expect that the regime will actually address their problems. As I argue in this chapter, this requires a democratic second-order framework.<sup>254</sup>

The conclusions of the previous chapters imply that we cannot reasonably expect any one specific mechanism of democracy—public deliberation, aggregation, selection, majority rule, etc.—to reach the quality of outcomes we want: decisions that in the long term and in general realize problem-solving.<sup>255</sup> Indeed, sometimes direct problem-solving may require the introduction of some intrinsically *undemocratic* elements within the political system as well. This is of course not as shocking as it sounds: market systems, for instance, are non-democratic, yet are of key importance in the coordination of collective behavior.

The main upshot of the whole experimentalism debate was of course that what works in terms of policy will only eventually be determined in practice. What this means institutionally is that no particular decision procedure can be expected to be good at first-order problem-solving across the decision spectrum: which procedure or institutional mechanism will be valuable for which problem areas is itself a question of considerable uncertainty. Hence, the institutional-design question itself has to be determined in practice (and through an experimental methodology) as well. Especially, as it may turn out, successful problem-solving may in some instances require the employment of non-democratic institutional mechanisms. However, as I show here, this pluralism of mechanisms of political decision-making has to be *embedded* within democratic structure to function adequately.

---

<sup>254</sup> For the concept of "second-order" democracy, see Adrian Vermeule, *Mechanisms of Democracy: Institutional Design Writ Small* (Oxford: Oxford University Press, 2007).

<sup>255</sup> They *may* of course do so, and certain particular instantiations of democracy *do* end up producing adequate outcomes. In the context of justification, however, we have to hold them to a higher standard.

This chapter therefore proceeds as follows: first, in sections 2 and 3 I show that the uncertainty assumption that we encountered in chapter 3 at the *policy* level actually applies also at the first-order decision-making level. In chapter 3 I called this the *Uncertainty Preservation Objection*. This implies that no single democratic mechanism can reasonably claim political authority in the robust way we need (that is, across a reasonably wide set of possible scenarios). Sections 2 and 3 respectively show that two standard models of democracy—the majoritarian/aggregative and the deliberative model, *cannot* be justified on the basis of Pragmatic Robust Instrumentalism. The problem with both is that any claims of reliability on which they depend rely on non-robust assumptions.

Second, section 4 presents another way of arguing for democratic legitimacy on the basis of the functional requirements of the experimental model of politics. The key argument is that an experimental form of democracy *robustly* fulfills those requirements, and its legitimacy can be based on that claim.

Finally, section 5 of this chapter suggests that the appropriate way to model democracy is instead as a *second-order* “control” mechanism that oversees the experimental first-order decision-making.<sup>256</sup> I construct the outlines of such a model in response to the functional requirements of experimental politics outlines in chapter 3 above. The upshot is that pragmatic robust instrumentalism calls for a specific type of *representative* democracy, and a specific understanding of the task of democratic institutions as exercising *control*.

Before I start on the argument, however, let me discuss some terminology. I use the term “a model of democracy” to denote a set of concrete mechanisms that describe

---

<sup>256</sup> See also Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*.

how the underlying value that are associated with the concept of democracy are realized. The idea of a model of democracy is therefore logically distinct from a normative theory of the value of democracy, such as the one described in part I. However, a model conception of democracy must have a connection to underlying democratic values. It must be internally coherent, but may choose to emphasize certain values or institutional realizations of democracy.

Therefore, we may distinguish a “direct” model of democracy, a “deliberative” one, a “participatory” one, maybe a “partisan” one: these are all bundles of (hopefully) coherent institutional mechanisms and assumptions about human behavior that describe how concretely certain democratic values should be exercised and realized. What does and what does not count as a conception of democracy is of course largely a semantic question, but there are presumably some boundaries, however fuzzy, beyond which a model of a political decision structure can no longer count as democratic.

As already mentioned briefly, the experimental pragmatic form of political legitimacy suggests that we should adopt a representative model of democracy in which the democratic function is exercised mainly as a second-order *oversight* or *control* function. Throughout this chapter I will call this understanding of democratic rule the *control model*. We can find expressions of this logic already in Bentham and James and John Stuart Mill’s ideas on representative democratic government. We can also find closely related ideas about democracy in later liberal thought, especially that of J. A. Hobson, and to some extent in the thought of John Dewey. Without going too much into it here, this model is characterized by a strict functional separation of the deputies (or representatives) and the electorate. In terms of the experimental model of politics, the role of the deputies



is to come up with a diversity of policy proposals that aim at resolving political conflicts and to implement them in an experimental way. The role of the electorate is—via appropriate institutional mechanisms—to ensure that deputies attempt to act in the public’s interest, and to provide the appropriate feedback to enable policy-adaptation. This latter role is exercised by issuing rewards (re-election) and punishment (removal from the legislature), and making explicit the approval or disapproval with the existing legislation.

This is of course not an entirely new idea of how democracy should work—basically as a variant of a principal-agent model with the people as the principal and the representatives as an agent. This model, in which representation is not a second-best approximation of a direct-democracy model, but has its own functional role, has recently received some renewed attention from a variety of sources: theorists of representative government,<sup>257</sup> republican thinkers,<sup>258</sup> and neo-pragmatists.<sup>259</sup> However, these thinkers tend to give markedly different reasons why a representative model of politics may be desirable: in Pettit’s case, for instance, in order to promote a robust conception of freedom. The consequence is that these conceptions emphasize very different aspects of representation, and also recommend very different concrete decision-making structures. The pragmatic problem-solving perspective gives yet another distinct view on this issue.

---

<sup>257</sup> Nadia Urbinati, *Representative Democracy : Principles and Genealogy* (Chicago, IL: University of Chicago Press, 2006); Bernard Manin, *The principles of representative government, Themes in the social sciences* (Cambridge: Cambridge University Press, 1997).

<sup>258</sup> Pettit, *On the People's Terms*.

<sup>259</sup> Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*; Kitcher, *Science in a Democratic Society*; Kitcher, “Public Knowledge and the Difficulties of Democracy;” Talisse, *A Pragmatist Philosophy of Democracy*.

Therefore, this chapter outlines in more detail the connection between the control model of politics and the pragmatic theory of legitimacy discussed in the first part. The difference between the control model and these other representative views is of course one of emphasis (and will become clear in the details). Just to highlight some of these differences in emphasis: in the control model, the main role of the deputies is not to represent *particular* interests (say, of their constituency, or of the social group they claim to represent). Accordingly, the role of elections is not primarily one of the selection of candidates that are most competent, closest to the people they represent, or are gyroscopically linked with their constituency.<sup>260</sup> Rather, it is the setting of the adaptive incentive-structure framework within which first-order decision-making takes place.

Interestingly, Dewey himself—surely the main figure linking democracy with pragmatism—in his own democratic theory did not always actually subscribe to a control model of democracy. Instead, he (mostly) favored a deliberative understanding. However, as I will show, he may have placed too much trust in people’s capacities (both those of the electorate and those of the prospective lawmakers). This again may be due to his underestimation of the depth of the uncertainty with which policy-makers are faced.

*First-order democracy and reliability claims*

However, before I get into the control model in more detail, I need to establish why other *first-order* models of democracy cannot be justified on the pragmatic conception of legitimacy. The general problem, as mentioned above, is that it is unlikely that we can

---

<sup>260</sup> For discussion of the ways to understand representation, see Jane Mansbridge, “Rethinking Representation,” *American Political Science Review* 97.4 (2003).

make a *robust competence claim* on the basis of those arguments. Recently, many so-called “wisdom of the crowds” arguments have been cited in support of claims of democracy’s first-order epistemic reliability.<sup>261</sup> The logic of those arguments tends to be—basically—that larger groups tend to make better decisions than smaller groups. These claims are based on the epistemic benefits of majoritarian judgment aggregation, or deliberative interaction. For practical reasons, it is of course infeasible to increase the size of decision-making bodies without limitation, but following the wisdom-of-the-crowds logic, we should at least approximate these mechanisms through the representative system. As I want to defend the control model of democratic politics, therefore, it is necessary to show why the “wisdom-of-the-crowds” model of politics does not fit with the assumptions of the pragmatic-experimental theory of political legitimacy. The reason is that these models are fundamentally premised on variants of the truth-tracking interpretation of universal reliability that I have already mentioned in chapter 3. Since truth-tracking arguments fail the robustness condition, the wisdom-of-the-crowds arguments are subject to the same criticism. This may not seem immediately obvious, as these arguments seem to be premised on the idea that some epistemic mechanisms can produce better epistemic outcomes than we could by ourselves. This point is based on what I have called the *Uncertainty Preservation Objection*, which, to repeat, is as follows:

In so far as the formal conditions under which a mechanism is expected to have a high truth-tracking ratio are defined at least partially with reference to the truth that is supposed to be tracked, the uncertainty about whether the conditions hold in a given situation preserves the initial uncertainty about the truth.

---

<sup>261</sup> Surowiecki, *The Wisdom of Crowds*; Landemore, *Democratic Reason*; Franz Dietrich and Kai Spiekermann, “Epistemic democracy with defensible premises”; List and Goodin, “Epistemic Democracy: Generalizing the Condorcet Jury Theorem;” Yann Allard-Tremblay, “The Epistemic Edge of Majority Voting Over Lottery Voting,” *Res Publica* 18.3 (2012): 207–223; David B. Hershman, “Two Epistemic Accounts of Democratic Legitimacy,” *Polity* 37.2 (2005): 216–234.

These “wisdom of crowds”-arguments tend to fall into one or both of two categories: the *aggregation* of disparate pieces knowledge (or “wisdom”) that amounts to better knowledge overall, or the *selection* of the correct decision from a pool of diverse pieces of knowledge. We can see how this would work: if the epistemic advantage lies in aggregation, then the more knowledge to be aggregated, the better; and if the epistemic advantage lies in selection, then the more diverse viewpoints, the higher the possibility that the correct one is among them.

In the following, I first look at two mechanisms that aggregate knowledge through majoritarianism: The Condorcet Jury Theorem (CJT), the so-called “miracle of aggregation.” After that, I look at Scott Page’s Diversity-Trumps-Ability Theorem (DTA) and its application to democratic politics by Hélène Landemore as an example of a deliberative mechanism of democratic policy-making. In particular I focus on how the logic of the theorems might lead to the conclusion that the more people involved in a decision-making process, the better.

The perplexing thing about the CJT and DTA is of course their compelling logic. As mathematical theorems, they *must* indeed hold whenever their assumptions are realized. Therefore, because CJT and DTA are axiomatic formulations we have to look at the assumptions they make, and crucially, whether and when those assumptions are satisfied in the “wild.”<sup>262</sup> In particular, the specific assumptions about the competence of decision-makers turn out to be crucial for the applicability of the theorems. It should not be surprising if there are some political situations in which more people do take better

---

<sup>262</sup> The wild world of politics, that is.

decisions. However, as already discussed in chapter 3 above, a successful instrumental defence of democracy under conditions of uncertainty requires us to demonstrate reliability whatever a good decision may turn out to be, or equivalently, if you like, demonstrate reliability to people who reasonably disagree about precisely what reliability actually consists of.

## 2. Aggregative/Majoritarian Models of Democracy

### *The Condorcet Jury Theorem*

So can we expect wide aggregation of votes to harness the “wisdom of the crowd” for the purposes of problem-solving? The Condorcet Jury Theorem (CJT) is the quintessential aggregative model of political decision-making, and it is most clearly concerned with the independently defined quality of its outcomes. For that reason, in this section I will discuss the CJT (in its most recent incarnation) as the key example of a model of an aggregative mechanism that grounds majoritarian models of democratic decision-making.

The theorem states that, under certain conditions, the larger the membership of a collective decision-making body, the more likely the collective absolute majority decision is to be the right one. The conditions are first, that each member, on average, has to have a chance of getting the right answer that is higher than pure chance; and second, that the votes members give are independent of each other. Strikingly, as List demonstrates, the CJT outclasses experts as soon as the membership reaches a relatively low threshold. That is, as we increase the membership, the effect of having a greater chance of getting it right

declines steeply in importance. Effectively, a large number of barely informed people (people with a, say, .51 chance of getting it right) may do better than a smaller group of experts (people who, say, get it right in 7 out of 10 cases).<sup>263</sup>

The conditions of the CJT in its basic form are:

- (1) If a group makes a decision by majority voting over two alternatives, and if
- (2) Every member of the group has a chance of picking the correct option that is greater than .5, and if
- (3) Every vote is statistically independent of all and every other vote (meaning that any information an observer might have about how some people in the group vote gives him no information about how the others vote), then
- (4) as the group size increases, the probability that the group will make the correct decision, approaches 1. If the conditions (1) and (2) hold, more decision-makers are always better than fewer.

Politically speaking, the remarkable result is that marginal competence of the decision-makers (provided (1) is fulfilled) matters not at all. Once the group is big enough, a majority vote will determine the correct answer with near certainty, whatever the competence of the members in it (as long as it is better than chance). This is good news, for instance, for people who question the ability of voters to actually select the *best* candidates for, say, parliament (as “elite” theories of democracies suppose is its real advantage). Even cynical observers, who doubt that voters manage to pick the best, might concede that the general public can at least pick candidates that are at least better than

---

<sup>263</sup> List and Goodin, “Epistemic Democracy: Generalizing the Condorcet Jury Theorem.”

chance. If decision-makers are mediocre, but generally ok, the result will be just as good as if decision-makers are a brilliant elite.

Various authors have extended the theorem in two interesting ways, strengthening its applicability considerably: firstly, List and Goodin have shown that even if there are more than two options, a large group operating with a plurality voting rule will pick the correct decision under the amended condition that<sup>264</sup>

- (1a) For every member of the group, the probability that she will pick the correct decision is higher than the probability that she will pick any *one other option*.<sup>265</sup>

Secondly, Grofman, Owen and Feld have shown that the group does not need to be homogeneous, as long as the first condition of the CJT is fulfilled *on average*.<sup>266</sup> Thus, condition 1 becomes

- (1b) For the *average* member of the group, the probability that she will pick the correct decision is higher than the probability that she will pick any one other option.

The result is that if conditions (1b) and (2) are fulfilled, the group will make the correct decision with a probability that rises with the group size. The striking thing is that therefore individual competence can be replaced by numbers: even for people who are

---

<sup>264</sup> *ibid.*

<sup>265</sup> This means that if there are for instance options A, B, C and D, and A is in fact correct, then for every member of the group, the probability for her picking A must be larger than the respective probability for B, C or D. The probability for picking A however can be smaller than the conjunctive probability for “B or C” or “B, C or D.”

<sup>266</sup> Bernard Grofman, Guillermo Owen, and Scott L. Feld, “Thirteen Theorems in Search of Truth,” *Theory and Decision* 15.3 (1983): 261–278.

barely better than chance at getting the question right, people who have, say, a .51 chance of getting it right on a dichotomous choice, we just have to find enough of them, and the group will make the right decision with near certainty. *And* adding even more people does no harm, as long as the conditions remain fulfilled.

Initially, this appears to be a very persuasive argument, because the conditions seem hardly stringent. Surely, many people will be better than a random process at getting a question right, and surely one could ensure statistical independence of votes. It seems therefore that the CJT can be the basis for just the robust *universal reliability argument* we need. However, with respect to real political problems this might be less straightforward.

The minimum competence condition for the CJT states that: for  $k$  possible choices, the probability that the average decision-maker will vote for the correct decision,  $k_i$ , must be larger than the probability to vote for any single other (wrong) option  $k_j$ ,  $i \neq j$ . This means that when the average decision-maker is more or equally likely to vote for a wrong option than for the correct one, the condition is not fulfilled. In such a situation, if the voters are on average really just guessing, as the group size and the number of options  $k$  increases, the likelihood of the group making the right decision approaches zero. The more open the political option space, the less applicable the CJT to real political contexts.

Enlarging decision-group membership therefore is a knife-edge issue: within a large enough group, average competence of .51 will ensure the correct answer with near-certainty, while average competence of .49 will ensure the wrong answer with equal near-certainty. Both average competences are, of course, for an outside observer for all



practical purposes indistinguishable from chance guessing. With large groups, the Condorcet majority vote process will deliver (depending on the actual competence of voters), either the right or the wrong answer with near-certainty.

The difficulty is that it seems that we have no good reason to believe that the competence condition holds with respect to a given question, unless we know the answer to that question. If we did, of course, we could dispense with the voting altogether. We need to have an *independent* reason (that is, independent of knowledge of the answer to a given question) to believe that the decision-makers will be better than random at finding solutions that work. What could that be? It seems almost obvious that (more or less) informed humans would have a higher-than-random chance of getting the decision right, that they would do better than a coin-flip or some other random procedure. So where might be cases where this does not hold?

There are two general ways in which the competence condition might be violated, and the CJT therefore would fail: first, if people are as good as chance (e.g. if they are truly randomly guessing) but the number of options increases towards infinity, the probability that they will get it wrong approaches certainty; and second, for a *finite* number of options, if people are *worse than chance*, the probability that the outcome will be wrong approaches certainty.

So let me assume for now that people are never worse than chance, but their competence may be equal to chance. In cases of unbounded policy space (where there are an indefinite number of possible policy options), or in cases of a bounded policy space that is infinitely divisible, this may arise. Consider an example of the latter: government

wants to optimize their tax revenue. On the one side of the government are followers of neo-Classical economics who believe that low tax rates will lead to higher revenue. On the other side are non-materialists who believe that people will work productively even under high tax rates. Assume also for the moment that there is a “correct” answer to this question. So what the government is looking for is the optimum point on the Laffer Curve. Let’s say this point happens to be at 32%. Let’s assume the government are Condorcetians, and therefore they throw the decision open to the whole population. Now people are no economists (notwithstanding that economists hardly have conclusively decided this issue), so let’s say they give a purely chance guess between 0% and 100%, each percentage point with equal possibility. Now the chance that a plurality vote will hit the right percentage, given that they are guessing only whole numbers, is minute. On an issue with so many possibly correct answers, the probability that a CJT-crowd will get it right is almost zero. If the crowd on average is as good a chance, but there are hundreds of possible options, the probability that the group will pick the right one will be approaching zero.

Now this example may be dismissed: Maybe the point is not to hit exactly on 32%, maybe possible options are between a narrow range of possible tax rates. Indeed, since we are not actually aiming for optimization, maybe we only have to make a choice between raising and lowering tax rates in a general sense. And if we look at political problems in this way, perhaps this issue actually does not occur in reality.

However, the more general problem this example identifies is this: the policy space (or rather the “policy solution space”) does not necessarily divide itself into a set of discrete options. Now, a crowd may be better than chance at choosing one from the

options presented to them—the competence condition may seem very plausible especially in binary cases, such as the context of a jury trial (which motivated the theorem in the first place). However, how can we be sure that the majoritarian judgment aggregation will choose “what works” out of the whole set of possible solutions? How can we be sure that the solution is even included in the set of options that is presented to the group? For example, can we be in any way certain that the average competence of a large group of decision-makers will pick an adequate solution to “solve the Eurozone debt crisis?” The point is that there is no clear set of possible responses to this question—it does not come pre-divided. For that reason the assumption of average consequence over an unbounded policy space is highly problematic.

This problem is exacerbated in cases of complex problems, i.e. those political problems that require a judgment on more than one decision. The solutions to complex problems are likely to require more than a single decision, but a conjunction of separate correct decisions. As List and Pettit point out, even if the group has a better than chance likelihood of making the right choice on each of the decisions, this means that the likelihood that they will get it right on the *conjunction* may be worse.<sup>267</sup>

Assume for instance this very simplified example: we have to decide whether or not to bail out the banks in response to the banking crisis of 2008. Now a bailout will be successful if and only if two conditions hold: the banks start lending again, *and* the government guarantees will not drive up the interest rate on government bonds. Getting the decision right, therefore, depends on our judgment whether the conjunction of those conditions holds. Now assume that our crowd can be assumed to make the right

---

<sup>267</sup> List and Pettit, *Group Agency : the Possibility, Design, and Status of Corporate Agents* 92-95.

judgment on the lending question with the average likelihood .6, and the same for the interest rate question.<sup>268</sup> Now if we have our group vote on whether to bail out the banks or not, they will actually only have an average .36 likelihood of accepting the bailout. As per the CJT, therefore, the group will almost certainly reject the bailout—whether or not it would actually be the right thing to do. Had they voted separately on the two conditions: the lending and the interest rate question, they would have been virtually certain to get it right.

What this illustrates is once again that the policy solution space does not come pre-divided. In those complex cases, *even if* the group would be reliable with respect to the individual elements of the problem, we cannot be sure that they will make the right decision with respect to the overall judgment on the whole issue.

There is a related issue with the miracle of aggregation mechanism as well: we might think that we *can* assume that the crowd is on average better than chance, on the basis that in a crowd that is mostly randomly guessing there are some people who have slightly better than chance idea on hitting the right answer within an unbounded policy space. Everybody else is guessing randomly. Then on average they will be more likely to choose the right percentage than any single other wrong one. In this case, *given that* the rest of the population votes such that none of the wrong choices have a higher probability than the correct one (this is satisfied when they vote perfectly randomly across a spectrum that includes the correct answer), the correct answer will edge out the incorrect ones by some margin. However, this margin increases as the number of good economists in the decision-making body increases, not as the total number increases (!). That is, if we know

---

<sup>268</sup> These probabilities are of course much higher than even necessary for the CJT.

that there are some people better than chance in a crowd that otherwise guesses randomly, we can increase the accuracy of the outcome by *restricting the membership of the group*. Especially at lower group number levels, this is a real possibility. The group may be better than chance, but worse than its best members.<sup>269</sup>

Thus, for some extremely complex problem like the ones mentioned, *either* we don't know whether a crowd is better than average, which means that we have no reason to trust the outcome, *or* we must assume that some people in the crowd are better than average, which means we should restrict the decision-making to only those people.<sup>270</sup>

Now, however, we have to look at the assumption that people are at least as good as a random guess. This does seem like a pretty unproblematic assumption. How could people be worse in their judgments than a coin flip?

For example, it would not be the case if the errors of the crowd are *correlated* in some sense. There are two ways in which this could happen: *cognitive* issues, and problems of *misleading evidence*. If some cognitive biases<sup>271</sup> are widespread among the population, this might violate the competence condition (and also the independence condition, possibly).<sup>272</sup> For instance, assume we are once again trying to determine whether we should raise or lower taxes to get the government deficit under control.

---

<sup>269</sup> *ibid.*, 95-97.

<sup>270</sup> There is of course another possibility: that we know the crowd is better than chance on average, but we do not know who in the crowd is raising the average score. This is of course theoretical possibility, but I would argue in this case we do not really *know* that the crowd is better than chance.

<sup>271</sup> I use “cognitive biases” here and in the following as shorthand for all cognitive processes within individuals that systematically (not randomly) distort how these individuals perceive facts and their environment.

<sup>272</sup> Kahneman, Slovic, and Tversky, *Judgment Under Uncertainty : Heuristics and Biases*; Kahneman, *Thinking, Fast and Slow*; Kelly, *Framing democracy : a behavioral approach to democratic theory*. Questionable as it may

How could cognitive biases lead to an average competence on this question lower than chance? Assume for instance that people always overestimate their own social class position, and therefore overestimate their burden of taxes. They may also vividly dislike paying taxes themselves, even though it would be in their net interest to increase the tax rate overall. In those cases, people may systematically err on the side of lower taxes. This may drive their average competence below chance. There is no room in this chapter to go into detail as to the cognitive biases, of which there is a rich literature, but it seems obvious that *if* there are cognitive biases that are widespread, and which therefore mean that errors are correlated, we have no reason to trust a mass decision. Ideally, if possible, we want to restrict decision-making to people who do not suffer from these cognitive biases, or at least those who are aware of them and try to correct for them.

A second way in which people's votes might be distorted to worse than random is through *misleading evidence*. If people base their CJT vote on their assessment of the general facts, but if some of these facts entering the decision are systematically wrong, the outcome is not likely to be the correct one. The misleading evidence could be a consequence of a deliberate action on some groups to influence the vote, or it could be the consequence of a common error (an erroneous factual belief that is, for innocent reasons, widely held). A quite good example of deliberately misleading evidence is of course the partly fabricated documents used by the US and UK governments to support the case for an invasion in Iraq in 2003. If the decision had been up to a crowd decision-making process (having access only to these documents), the likelihood that the risk would

---

otherwise be, see also the argument in Bryan Caplan, *The myth of the rational voter : why democracies choose bad policies* (Princeton: Princeton University Press, 2007)

have been assessed correctly is quite small. The truth about the effect of tobacco smoke on people's health prior to the 1980s may be a case of a erroneous, but widely held belief—if the science at the time simply was unable to determine health risks of tobacco.<sup>273</sup> Substance control laws and lack of restrictions on smoking in public places might well have been influenced by this particular piece of misleading evidence.

In any case, we might conclude that in cases where bias or error are likely, we should not use a Condorcetian process to determine the correct answer. Consequently, in those cases we ought not to expect correct answers, and the basis for an epistemic argument for democratic legitimacy becomes a lot narrower.

The problem for the application of the CJT to a justification of democratic authority may be even more severe: now we have to distinguish how we understand the misleading evidence in question, which itself depends on how we understand the CJT and its role in legitimating democracy. Consider the following argument, made by Dietrich in an important paper.<sup>274</sup> On the one hand we may think that for *any particular political decision*, the CJT means that the decision is likely to be right. On the other hand, we might think that *in general*, that is, over many decisions, the CJT gives us reason to believe that decisions tend to be correct.

If we take the former position, then we may conclude that any misleading evidence in the particular vote is only a part of the complete set of common environmental factors influencing all of the voters in the same way—and the votes can

---

<sup>273</sup> Although some people may think that this also was deliberately misleading evidence, in this case the cigarette companies and their research institutes being the misleading party.

<sup>274</sup> Dietrich, "The Premises of Condorcet's Jury Theorem Are Not Simultaneously Justified."

still be independent, *given* the circumstances. However, in this case, the competence condition, condition (1) of the CJT is likely to be violated: if misleading evidence is part of the common evidence, people are likely to be worse than random at getting it right—or at least we cannot as a rule assume that people are better than random in the particular case.

If we take the second position on the CJT, namely that it holds *in general*, then it might be more plausible to assume that the competence condition holds (that decisionmakers are *overall* pretty good). However, in this case we cannot argue that *given* the misleading evidence is common evidence, the independence condition is satisfied. If we want to claim that the CJT method has a *general tendency* to get things right, we have to conceive of the misleading evidence as a *variable factor*, extraneous to the structure of the problem. And in that case, misleading evidence would violate the independence condition. An analogous argument can be made for cognitive biases.

The consequence of this is that in the presence of misleading evidence or cognitive biases, the conditions of the CJT are unlikely to be fulfilled simultaneously. To illustrate, assume that we are trying to assess whether a mass vote by the general public is a good way to decide whether the risk another country poses is worth going to war.

We can either look at a particular instance, say whether to go to war with Iraq in 2003 or not. Then we can guarantee independence of votes when we assume that the fabricated evidence about Iraq's threat is part of the given circumstances. Everyone has access to the same evidence, and may still vote independently. Of course even though



independence is guaranteed, the fabricated evidence means that the competence condition is likely to be violated.

On the other hand, we may look at how the CJT method performs *in general* on decisions whether to go to war. Then it may be justified to assume that people get it right on average. However, then in our particular case of the Iraq war the presence of the misleading fabricated evidence cannot be part of the given common factors, and it means that independence is violated.

What does this mean concretely? It means that unless we can guarantee the absence of cognitive biases and misleading evidence, we cannot assume that the CJT's conditions are simultaneously justified, and we cannot assume that larger groups will in fact make better decisions. However, cognitive biases and misleading evidence are defined with reference to the *correct solution*—a cognitive bias is a mechanism that distorts the probability that an individual will *find the right answer*, and misleading evidence is evidence that suggests the correctness of some *false* answer. Therefore, we cannot in general assume that they are absent—just as we cannot assume that they are present. Hence such assumptions violate the *robustness* condition—either assumption could be reasonable rejected, on the basis of reasonable disagreement.

#### *Independent grounds for assuming competence*

But may there not be some independent reasons to trust that these conditions obtain, i.e. some reason to believe the competence condition should hold, without having to define some standard of rightness? I have already briefly considered Dietrich's argument that it

would be irrational consistently to guess worse than chance in the section on historical arguments in support of universal reliability in chapter 3 above. However, consider the following argument:

- (1) For any given  $P$ , if people sincerely disagree about the truth-value of  $P$ , and if they are at least vague epistemic peers,<sup>275</sup> one side must be in error.
- (2) Since it is less likely for an individual to be in error than it is to be right, it is generally *more* reasonable to assume that the minority is mistaken about  $P$  than the majority.
- (3) Therefore, it is more reasonable to assume that the competence condition of the CJT holds than the reverse.<sup>276</sup>

This argument turns on the question whether it is reasonable to assume premise (2) given ignorance about which decisions are right and which are mistakes.

Now it is of course difficult to make such a judgment about what is reasonable and what is not. However, remember that in order to satisfy the conditions of PRI, we would have to make a *robust* case—that reasonable people could accept—in support of (2). The most significant problem to make such an assumption becomes clear when we look at political and moral activity dynamically, or rather, *historically*.

---

<sup>275</sup> This means that roughly speaking, they have the same evidence, and roughly speaking, they have the same cognitive capabilities, and that they recognize this fact.

<sup>276</sup> This basic argument has been made in various forms by Hershenov, “Two Epistemic Accounts of Democratic Legitimacy;” Allard-Tremblay, “The Epistemic Edge of Majority Voting Over Lottery Voting;” See also the discussion in Robert E Goodin and David M Estlund, “The Persuasiveness of Democratic Majorities,” *Politics Philosophy Economics* 3.2 (2004): 131–142.

First, then, it seems to me that for factual or causal beliefs, (2) obviously does not necessarily hold: quite clearly, if there is reasonable disagreement about the causal setup of the world, we cannot just assume that any minority is more likely to be mistaken than the majority. In science as well as in politics onetime minorities turn out to have been correct all along with respect to causal beliefs about the world. Consider the germ theory of disease, whose proponents (especially the Hungarian physician Ignaz Semmelweis) were considered at best harmless cranks and at worst destructive enemies by the majority of the medical community.<sup>277</sup> Examples of this sort abound, and the frequency of reversals of majority opinion with respect to these suggest that it is not a reliable guide to truth. Another example: before the start of the Iraq war, 55 per cent of Americans believed that Iraq possessed and had hidden away weapons of mass destruction, falsely as it turned out.<sup>278</sup> Given the ubiquity of misleading reports or distorted evidence of this kind, we cannot say in a given situation that the majority is right unless we know *independently* what the right answer is. Note, however, that the evidence was misleading *only because* it did not turn out to be true that Iraq had WMDs.

A look at the history of ethical (and political) progress suggests that the same problem arises with respect to normative questions as well. Quite frequently, what we now consider the clearly morally superior view has been a minority view for centuries, be it on slavery, religious toleration, or women's rights. Again the problem would be, how to explain reversals of majority opinion. Take for instance the question of the death penalty. To save the thesis we would have to endorse the moral relativist view that the moral

---

<sup>277</sup> Carl G Hempel, *Philosophy of Natural Science* (Englewood Cliffs, NJ: Prentice-Hall, 1966) 5-6.

<sup>278</sup> Gallup Poll, June 16, 2003. <http://www.gallup.com/poll/8623/americans-still-think-iraq-had-weapons-mass-destruction-before-war.aspx> ; accessed April 2013

rightness/wrongness of the death penalty also reversed right around the time when majority opinion changed (but not *because* of it<sup>279</sup>), *or* the claim that the reversal in majority opinion *is explained* by a change in institutional circumstances. Neither of these is particularly compelling, seeming, as they do, somewhat *ad hoc*.

Incidentally, the same issue arises with respect to other aggregative mechanisms whose conditions are defined with respect to the correct answer. Consider the so-called “miracle of aggregation,” often illustrated with Francis Galton’s famous case that the median estimate of a crowd of the weight of a particular dressed ox was more accurate than any of the individual estimates.<sup>280</sup> The reason is that the guessing errors in this case were uncorrelated, and clustered randomly around the true value. The people who guessed too high and those who guessed too low “cancelled out.” When this occurs, the mean or median guess will be more accurate than that of a randomly drawn individual from the crowd. The true signal, as it were, is buried in random noise.<sup>281</sup>

Is it reasonable to assume that in political decision-making errors are uncorrelated and randomly distributed about the mean? In that case, we have to rule out systematic biases and (again) misleading information. This is the same problem already discussed above: For the miracle of aggregation, we do not have to know *what* the truth is, but we

---

<sup>279</sup> Otherwise the argument invalidate the procedure-independence of truth required for an epistemic justification.

<sup>280</sup> Note that in the original report of the case, the median was most accurate, not the mean. This is contrary to more recent versions. Note also that in that case there must have been an even number of guesses, otherwise the median would have coincided with at least one individual guess.

<sup>281</sup> Rousseau’s theory of the discovery of the general will is also often interpreted in those terms. Another well-known statement of the miracle of aggregation can be found in Benjamin I Page and Robert Y Shapiro, *The rational public : fifty years of trends in Americans' policy preferences, American politics and political economy series* (Chicago: University of Chicago Press, 1992).

do need to know independently what the *relation* between the truth and everyone's opinion is.

These historical considerations I have just sketched therefore pose something of a problem for these types of argument: unless these cases can be explained away, the assumption that as it happens, in our current situation the minority is on the whole more likely to be in error than the majority is actually not that intuitive. This is not to say that the opposite assumption—that the majority is more likely to be wrong—is any *more* plausible. If it were, majoritarianism would turn out a terrible form of government. It would almost always guarantee a wrong decision. Rather, the point is that because we need to have access to the truth in order to decide this question *either way* with confidence, we cannot claim reliability for democracy without it.

This means that we cannot claim that any particular majoritarian-aggregative procedure is likely to get the right solution *unless* we want to assert some independent standard for the problem-solving on that claim is based. But that would violate the robustness condition.

### 3. Deliberative Models of Democracy

But perhaps universal reliability can be found in a different democratic mechanism: the *public deliberation* about policy proposals. The public competition inherent in democratic decision-making means that political beliefs have to be defended with proper arguments. Through deliberative contestation—the “forceless force of the better argument,” as

Habermas' famous formulation has it—everyone's mistaken assumptions are challenged and hopefully revised in favor of better-supported beliefs. In the ideal case, the best argument would eventually convince everyone and there would be consensus. But even in a case where consensus cannot be reached, people would have better beliefs post-deliberation than before—maybe even a majority would have converged on the best argument—which means that a vote post-deliberation would be more likely to be correct.

*The Assumptions of Deliberative Models*

The first issue is of course that this requires a number of assumptions about people's motivations and cognitive capabilities. In short, everyone involved in the deliberation must be focused on solving the problem at hand, sincere in their arguments, and be willing to revise their own beliefs in the light of others' arguments. The “pathologies” of real-world deliberation that suggest otherwise are well documented.<sup>282</sup> Besides the obvious point that the group realistically might well converge on some non-optimal argument (for instance the most intuitive, the one that can most easily be described in a coherent narrative, or the one proffered by the most rhetorically gifted), there is another problem: the *best* argument might not actually be *right*. There could be convergence or even consensus on the most cogent argument, but it would still be wrong. The deliberation argument requires that the following *correspondence condition* hold:

If policy *P* is most reasonable within the deliberating population, given the circumstances, it is also right, and if policy *P* is right, it is also the most reasonable, given the circumstances.

---

<sup>282</sup> See for instance, Sunstein, *Going to Extremes : How Like Minds Unite and Divide*.

This condition fails to hold when the correct solution appears to be completely unlikely. As mentioned already in the previous section, we have reason to resist the assumption that what appears to be a good reason to most people is also the right reason. Through most of history, denying women higher education was justified based on the “obvious” fact that women were not suited to the intellectual or professional life and that their temperament would mean that education was wasted.<sup>283</sup> I take it that it is obvious to any reasonable person now that the former view was wrong. However, before the contemporary view gained traction, consensus on the obvious would not have yielded the correct view—indeed it took more than a century of activism by feminist moral entrepreneurs for the views of what is reasonable to change. Even if people’s unreflective views were challenged in deliberation, there was no reason to reject such an “obvious” premise simply because a small minority held it to be wrong. This is not to say that deliberation has not helped in disseminating the new, better ideas: it probably has. However, from the mere occurrence of deliberation we cannot infer whether the correspondence condition holds in that context.

An historical perspective again illustrates the problem: what are we to make of situations where there was an deliberation before and after a change in collective moral beliefs, and therefore the change in beliefs cannot be explained by the presence or absence of deliberation? Surely, for instance, there was some degree of deliberation in the British parliament before, during, and after the abolition of the slave trade in 1807 and we have no immediate reason to assume that institutional conditions shifted precisely at this time. In general, we may ask: does democratization engender “better” moral beliefs,

---

<sup>283</sup> Kitcher, *The Ethical Project* 145-153.

or the other way around? This suggests that the occurrence of deliberation alone cannot be the sole explanation of the shift in moral beliefs that occurred around this time.<sup>284</sup> It seems therefore that deliberation will produce the right results only if enough people already have the right beliefs—or as we might put it, the right beliefs seem reasonable or cogent enough to most people that they let themselves be persuaded in deliberation. There must be a correspondence between rightness and “obviousness.” And of course, unless we know what the right beliefs are, we cannot really form an expectation of such correspondence. More generally, often one view and its reverse can seem equally reasonable. “Common sense” sayings often have an exact counterpart. Too many cooks spoil the broth, while many heads are better than one. As Paul Lazarsfeld put it at one point, “Obviously something is wrong with the entire argument of ‘obviousness.’”<sup>285</sup>

Again, therefore, the epistemic democrat seems to be forced into the uncomfortable position of having to either adopt an *ad hoc* moral relativist position (such that the “truth” about slavery changed in tandem with the beliefs about slavery), or to argue that somehow deliberation was not functioning properly at that time, which explains the collective moral failure—but then conversely also cannot explain the instance of what appears to be moral progress. The point once again is that the success or otherwise of deliberative practices is premised on specific conditions—and there may be reasonable disagreement about whether these conditions obtain, at any given time, or across a wider range of issues.

---

<sup>284</sup> This “shift” of course occurred over a long period of time, driven to a large extent by abolitionist activists.

<sup>285</sup> Paul F Lazarsfeld, “The American Soldier-An Expository Review,” *The Public Opinion Quarterly* 13.3 (1949): 380; cited also in Watts, *Everything Is Obvious: Once You Know the Answer*. Watts also cites the example of contradictory common-sense sayings.



This issue cannot be explained away by assuming more ideal forms and conditions of deliberation, short of postulating full information. The problem arises even with ideal deliberative behavior. Even if we grant that everybody honestly takes into account each other's arguments, deliberation will be able to play its epistemic role only if there is a correspondence between a "good reason" and the *truth* at a given point in time. This is not a given: if there is uncertainty about the truth, it is in principle reasonable to reject the correspondence condition.

Once again, to be sure, uncertainty about whether the competence or the correspondence conditions hold does not mean that they do not *in fact* hold in a given instance. Recall however that robustness requires the proposed mechanism to exhibit its epistemic advantages given a range of possible reasonable scenarios of cognitive competence and of the nature of political problems. The deliberation argument, for instance, holds only in a very narrow range of situations: whenever the proposal seeming most reasonable to most people corresponds to the correct one. It is therefore not robust in this sense. This would not be a problem if we could make a case that, generally, the correspondence condition holds. The argument of the last two sections has been that uncertainty about that truth re-appears here as uncertainty about whether the relevant conditions hold.

### *Diversity Trumps Ability*

However, there may be another way to show that under some conditions, deliberative procedures may be reasonably expected to reach the right solutions, regardless of what

those right solutions turn out to be. However as before, the issue with this argument is that it runs afoul of the robustness requirement, as I will outline below. I will, however, return to this argument in the first substantive section of the next chapter, showing how diversity can be utilized within a control model of democratic politics.<sup>286</sup>

Scott Page's Diversity-Trumps-Ability Theorem (hereafter DTA) states that, under certain conditions, the more cognitively diverse a group is, the better the decision will be. More precisely, this states that *given the same pool from which the decisionmakers are selected, a random diverse collection will outperform a selection of the ones with the highest ability*. "Ability" here is understood as basically scoring high on the *same* scale of measuring ability: that is, the best could be the ones with the highest IQ, or more relevantly, perhaps those most closely acquainted with the problem area at hand. This leads to a similar conclusion as the CJT, namely that the individual ability of decisionmakers does not matter much once there is sufficient diversity in the group.

Strictly speaking, this of course does not say that larger groups will be better than smaller groups, but under certain circumstances it also (implicitly) leads to the conclusion that the more people involved, the better the decision. How this is so I will point out in a minute. For the DTA to hold four conditions have to be fulfilled:<sup>287</sup>

- (1) The problem cannot be too easy, i.e. it cannot be the case that everyone by themselves will always solve the problem.

---

<sup>286</sup> The writer who probably most closely associates democratic legitimacy with the epistemic benefits of diversity is H el ene Landemore. See Landemore, *Democratic Reason*; H el ene Landemore, "Deliberation, Cognitive Diversity, and Democratic Inclusiveness: an Epistemic Argument for the Random Selection of Representatives," *Synthese* (2012); see also H el ene Landemore and Jon Elster, *Collective Wisdom: Principles and Mechanisms* (Cambridge: Cambridge University Press, 2012).

<sup>287</sup> Page, *The Difference* 158-62

- (2) Every member of the group will have a grasp of the structure of the problem. In other words, everyone in the group can distinguish a solution that is more likely to be true from one that is less likely to be true.<sup>288</sup>
- (3) The people have to be diverse in their outlooks and perspectives on the problem. In particular, it must be the case that if one member proclaims a solution that is not in fact the correct solution, there must be at least one other member who can point out at least one reason why it is not the best solution.
- (4) The pool from which the decision-making group is drawn is large, and the group drawn from it is also relatively large.

Now, the conditions leading to a “wisdom of crowds”-type conclusion are (3) and (4). More people increase the likelihood that (c) holds. Of course it is consistent with the DTA that small diverse groups may perform better than large uniform ones, but it is also implicit in the theorem—and is taken by interpreters to be implicit (e.g. Landemore)—that adding more people to the small diverse group will not make it worse. What the DTA boils down to in terms of institutional advice is this: when selecting a group of decision-makers, don’t restrict the group to the ones you think are best suited, but include as many diverse opinions as is practical. Indeed, I have no issue at all with this view, I think it is highly plausible. However, I doubt its applicability as a supporting reason in the context of an outcome-focused justification of democracy.

---

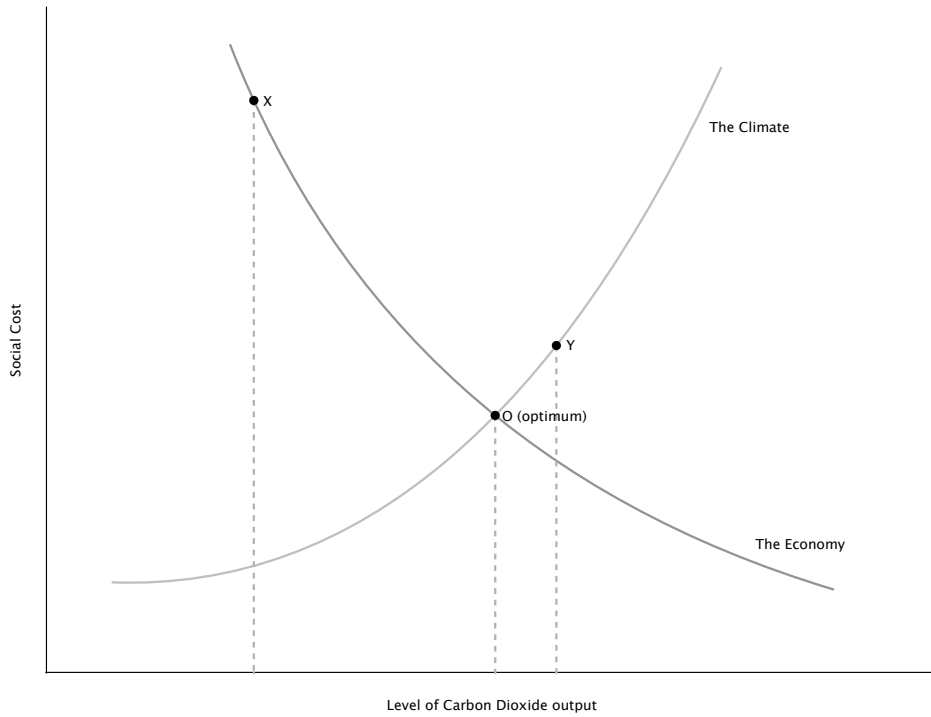
<sup>288</sup> In Page’s words, this condition means that everyone involved “has to know calculus.”

Now the DTA does not really state outright that having more people involved in a decision is necessarily better, and for that reason the connection to democratic decision-making may be somewhat tenuous. The condition that (to some extent) acts to limit the group size is of course condition (b), which basically states that adding too many people who have no clue may make the outcome worse. But *if* there could be some screening mechanism that restricts the pool of people to those having some idea about the problem, *ceteris paribus*, the DTA states that the bigger the group, the better. That is, the DTA allows for a significant restriction of the *pool* from which the membership of the decision-makers are drawn, however, from *within that pool*, the more people involved the better.

Now however, let us take a closer look at the conditions of the theorem. Condition (1) is relatively straightforward. It is obvious that for really easy problems nothing will outperform a single decisionmaker. Neither though will adding more decisionmakers in this context make the decision worse. As it were, if the task is merely boiling water, one cook will do just as well as five cooks or as well as two hundred amateurs collaborating. So for the purposes of this chapter we can exclude cases that fail condition (1).

Condition (2), which postulates some basic minimum level of competence, is more complicated. This basically states that people can distinguish a better argument from a worse one, and essentially also that they can identify the correct (or best) solution, when it is found, as such. In order to best understand this idea we should, as Page does, see complex epistemic problems as multidimensional estimation problems. Consider for instance a simplified topical example: Imagine a country called Hermetica that has a closed climate and a closed economy—that is, it's climate and economy are only affected by what the country's own CO<sub>2</sub>-emissions. Now the government wants to optimize the

level of CO<sub>2</sub>-emission of Hermetica's economy. Imagine the effect of the level of CO<sub>2</sub>-emissions is as in Figure 2.



*Figure 2: A Complex Estimation Problem*

In this case, the optimal solution is to reduce it so much that the marginal social benefit of reducing Hermetica's greenhouse effect caused by another unit reduction in CO<sub>2</sub> is equal to the marginal social cost of reducing CO<sub>2</sub>-emissions by another unit. This is a two-dimensional problem that requires estimating *both* the effect a unit reduction of

CO<sub>2</sub>-emissions has on the climate, and the adverse effect it has on the economic product.<sup>289</sup>

Now we know that there is an optimal point. This is the point O in Figure 2. The competence required by (2) is *not* that everyone should know this optimal point from the start (this would violate (1)). However, everyone should be able to identify a better proposal from a worse one.

Consider that two scientists of Hermetica's institutes of meteorology and political economy respectively are tasked with finding the optimum. The meteorologist knows only the shape of the "Climate" curve, therefore his optimum would be somewhere to the left of the graph, close to zero emissions. Say therefore the meteorologist advocates reducing the emissions sharply, to the point X. Now the economist points out the social cost of this reduction and suggests that the point Y might be better. Condition (2) requires that *both* recognize that Y is better than X, since it reduces the difference between the marginal cost and benefit, and is, as it were, closer to the optimal point O. They would not fulfill the competence condition if they could not conclude that Y is better than X (and that O is better yet). They would literally just not understand the problem, and would not know how to find a solution.

Condition (3) is the one that specifies the *kind* of cognitive diversity that is required for the DTA to work. The optimal solution to the problem, then, is one that is an optimum on *all* of the dimensions of estimation. This means that no-one, on their dimension of estimation, can find an improvement. Since by assumption (1) no single

---

<sup>289</sup> Hermetica, somewhat unrealistically, has an economy and a climate that are very easy to estimate. Therefore we treat the two individual estimates as simple ones, although they themselves are of course complex multidimensional estimations.

decision-maker will know the structure of all the dimensions, condition (3) merely ensures that in the collective, all of the dimensions are represented. Thus, for instance, two people by themselves could simultaneously fulfill conditions (1) and (3), although obviously, having more people increases the likelihood to satisfy in particular (3).

As an example, let's return to Hermetica. Condition (3) would not be fulfilled if there were only meteorologists in the room, who have no clue about the precise shape of the economy function. Maybe they are aware that there is such a thing as the economy, and perhaps they vaguely know that the economy curve would slope down from left to right. In that case, they may still proclaim solution X as the optimum. And nobody would point out to them that X does not efficiently balance the effects on the climate and the economy. It is this kind of diversity that the DTA requires. We can see how condition (3) ensures that ability is trumped by diversity. We can add the best of the best meteorologists to the decision-making process, those people who can absolutely precisely estimate the relationship between Carbon Dioxide and the Climate, without an economist in the room they will not find the optimal solution to the problem.

Note as well the difference between Page's competence condition and the CJT's competence condition. Here people only have to *understand the problem*, which can perhaps be ascertained without knowing the solution. In the CJT case, people have to *have a good chance of getting the problem right*, which cannot be ascertained without knowing the right answer. The DTA's competence condition is already more realistic than the one required by the CJT.

Finally, condition (4) ensures both that the random diverse group is in fact diverse, and that it is significantly more diverse than the group selected by ability. If we only draw a few people, of course, the group will be less diverse than if we draw more. This is the “wisdom of crowds” idea here. If however the pool from which to select were small relative to the people selected, the best could be just as diverse as the diverse group.

Now do these conditions really reflect political practice? I concentrate mainly on Page’s conditions (2) and (3), the minimum competence and diversity conditions. Cases in which condition (1) is violated are relatively uninteresting, and even if there are practical problems with condition (4), at least in theory it could easily be fulfilled. I shall therefore assume that political problems are hard, and that groups selected by ability are always less diverse than groups selected on their diversity.

I take a problem-based perspective, trying to show how conditions (2) and (3) may not be as plausible as they look. Whereas Page starts from a certain idea of how complex problems are structured, I shall argue that quite frequently, political problems do not fit the pattern, meaning the DTA cannot be applied. I look at four possible types of cases, all of which are common to the political world: problems where people are generally incompetent, problems for which it is difficult or impossible to ascertain a solution with certainty, problems which are embedded in a lot of noise, and problems which are difficult, but not *complex* in the sense of the DTA.

First, of course, there is the possibility of a straight violation of condition (2). If people have no idea of the basic structure of the problem at hand, we should not add more of them to the decision-making process. For instance, if people are unaware that



reducing CO<sub>2</sub>-emissions also has a social cost, as it tends to reduce economic output,<sup>290</sup> adding more of them to the process will not make finding the optimum any easier. Depending on the problem at hand, the minimum competence required can be actually quite high: it requires everyone involved to recognize the basic shape of the problem, and it implies that everyone involved has to understand how the arguments of the people work that have viewpoints different from one's own. The DTA requires you to both have a "maverick" perspective that is different from the ones of the supposed "experts," *as well as* a good understanding of the problem. For moderately complex problems like Page's examples drawn from business or marketing this might not be a difficulty. Consider however an extreme case, for instance a highly complex and specific problem in physics—one of those where an understanding of the problem already requires you to become an "insider." For many issues that affect policy in an important way, it might well be that in the pool of people understanding the problem in the sense of (b), there are only a handful of people.

Even if we concede that they are not always like scientific problems, political problems may often be rather complex—perhaps halfway between science and common sense. Then adding more people to the decision process might well fail to clear the bar of the competence condition (despite the presence of many self-declared "mavericks" in politics). Thus, the more complex an epistemic problem that comes up in politics is, the less likely it is that condition (b) is fulfilled.

Consider an example: Hermetica's primary education system, which the efficiency-focused government wants to optimize. In particular, what is to be improved is

---

<sup>290</sup> If they assume that the "Economy" curve in Figure 2 is a straight horizontal line, for instance.

the ability of children to cope with a rapidly changing world of employment later in life. Like everywhere else, the effects of instruction on children's brains are highly complex to grasp, and perhaps still not universally understood. In Page's terms, let's assume that it is a problem with many possible dimensions, not all of which are yet known. Now in Hermetica there is a (moral) majority who have no idea of educational science and childhood psychology but think discipline and corporal punishment alone are a great, traditional way of bringing up children. They know one dimension, and are unaware that there are others to the problem. Including these people in the decision process will violate condition (2), since they do not understand the optimization problem in the correct way.

Now a defender of DTA may argue then that we should exclude these people from the pool from which decisionmakers are selected and select a diverse bunch from the smaller pool. If however, the remaining pool that is available becomes smaller and smaller, condition (4) might be violated (and the principle seems less and less democratic). We can see thus, how with problems of extreme complexity, selecting the most diverse group might either lead to the inclusion of people who don't understand the problem, or force us to drastically restrict the pool of potential decisionmakers, until we end up with only a narrow group of experts.

Not only people who don't understand the structure of the problem can violate condition (2), however. Consider cases of *logical complexity*. if people know the basic structure of the problem, but are unaware of how the dimensions to be estimated are structured—if they don't know the general shapes of the curves in Figure 2, for instance—then adding more of them to the decision process will not improve the outcome. For instance, people may assume that the effect of CO<sub>2</sub> in the climate is close to linear,

whereas in reality there may be so-called “threshold effects” that have lead to disproportionate consequences once a certain threshold of concentration is reached.

Similarly, if the problem is such that not all dimensions are known, adding more people who believe otherwise to the decision process might make it worse. In other words, let’s say there is a third mechanism in Hermetica that links CO<sub>2</sub> to social cost,—for instance, reducing CO<sub>2</sub> emissions will cause the plant population to shrink—then a group that fulfils Page’s conditions, but lacks some biologists, may well miss a part of the problem.

However, this depends on the new perspective actually making an obvious improvement, actually getting the group closer to the optimal solution—and this has to be obvious to everyone involved (as per condition (2)). The new dimension has to be, as it were, a “Eureka”-solution.<sup>291</sup> That is, a solution that can be immediately verified because it explains some missing dimension or because it “fits” into the structure of the unsolved problem like the missing word in a crossword puzzle. But the kinds of political issues we are looking at here are not necessarily like that: such confirmation ex ante is not always, indeed probably only rarely, available.

The biologist could propose the effect of CO<sub>2</sub>-reduction on the plant population, or the meteorologist could warn of further risks. To some extent of course it could be tested if these effects have been present in past climatic events, but not necessarily so, given that these proposed effects are rare and (by definition) not obvious. Thus, even if

---

<sup>291</sup> Cass R Sunstein, *Infotopia: How Many Minds Produce Knowledge: How Many Minds Produce Knowledge* (Oxford: Oxford University Press, 2006) 60-64.

the true missing dimension is proposed, it might not be clear that it is the true missing dimension—especially if the correct one has to be picked out from a number of proposals.

What is required therefore in such a situation is that the decisionmakers are aware of essential uncertainties, the possibility of hidden dimensions, and have a grasp of their plausibility—in other words, a “scientific mindset.” I would think that this characteristic is both *not* universal, *and* can be relatively easily ascertained in individuals. And consequently, the decision can be improved by restricting the pool of potential decisionmakers to these “experts.”<sup>292</sup> This is another way in which expanding the number of decisionmakers might, in the end, be violating the competence condition.

#### *Other kinds of problems*

There is, however, another category of possible problems: there might be relatively simple estimation problems that are nevertheless embedded in a lot of “noise,” i.e. actually *irrelevant* factors which nevertheless seem they might be relevant. The problem is one of “overfitting” one’s theory to all the seemingly relevant, but actually spurious, factors.<sup>293</sup>

Here the necessary competence would be the ability to identify the correct solution while disregarding the irrelevant dimensions. Consider again Hermetica’s CO<sub>2</sub>-output. Given that it’s a fictional example, its causal relationship between CO<sub>2</sub>-emissions and social cost is two-dimensional and remarkably straightforward. However, in reality

---

<sup>292</sup> Of course, *within* this pool one could still try to get as much diversity as possible.

<sup>293</sup> See for instance, Gerd Gigerenzer and Peter M Todd, *Simple Heuristics That Make Us Smart* (New York: Oxford University Press, 1999); Gerd Gigerenzer, *Rationality for Mortals: How People Cope with Uncertainty*, Evolution and Cognition Series (New York: Oxford University Press, 2008).

there are dozens of other possible candidates for a curve in the graph, some of which may have a genuine effect, and some of which may be spurious. As in the example above, the plant population could be one of them, or the effect on the coal mining industry, and so on. Opening the decision process to a diversity of individuals or groups in a consociational decision process would probably amplify these irrelevant factors as everyone contributes their particular perspective. Again, of course, if for every proposal one could easily verify whether it is relevant or not, this would not be a problem. However, with political problems, as argued above, this is often not the case.

But can we really know whether we are faced with one of these, as Gigerenzer calls them, “*less-is-more* problems”, that is, problems where less information actually facilitates better decisions?<sup>294</sup> For instance, we might look at policy areas where high-information methods have a poor track-record, especially compared to simpler ones. Gigerenzer’s own example is the stock market, where the apparent informational advantage of insiders does not necessarily translate into a decisional advantage when compared to amateurs.<sup>295</sup> Many issues of prediction in actual political practice arguably have a similarly weak track record.<sup>296</sup> However, in the end we cannot reliably say whether a problem is complicated or merely embedded in a lot of noise. The lesson to be drawn is that we probably won’t know with what we are going to be faced.

The final set of cases I want to discuss are those where the difficulty is not a consequence of the complexity assumed in the DTA model. For instance, consider the

---

<sup>294</sup> *ibid.*

<sup>295</sup> See also Kahneman, *Thinking, Fast and Slow*.

<sup>296</sup> The authoritative study here is probably still Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?*

possibility that the problem itself is not a multidimensional estimation problem at all. It might be either a simple *single-dimensional* problem, a problem whose difficulty derives from its *computational complexity* or from the fact that it requires difficult-to-obtain *facts*. In those cases, diversity does not add any benefits. This does not mean that these problems are *easy*, i.e. it is not the case that condition (1) would be violated, it is only that their difficulty is such that diversity, and the wisdom of crowds does not help.

For instance, assume the problem is not a logically complex one like the effect of CO<sub>2</sub>-reduction on social cost, but one merely of aggregating information in a single dimension. Take as an example the problem how to most efficiently reduce energy consumption overall—this is an important problem both for individuals at home, and for governments that want to design public policy to incentivise or disincentivise certain behaviours. The problem is simple since it involves only one dimension to be estimated (kilowatt hours), which is even easy to measure, but the problem is of course still complex since it involved adding and comparing the energy use of myriads different places of energy consumption. Even within a home, the information is spread very thinly. This problem therefore calls for plenty of *computational power*, rather than a diversity of viewpoints. In fact, it seems, the general public are remarkable bad at estimating the energy consumption of even their own home, as a recent report argues.<sup>297</sup> The question of how to reduce energy consumption is *logically* simple, but *computationally* complex. As such its solution does not require many different perspectives.

---

<sup>297</sup> Shahzeen Z Attari et al., “Public perceptions of energy consumption and savings,” *Proceedings of the National Academy of Sciences of the United States* (2010) In general people seem to overstate the effect of free and/or easy-to-do measures, like switching off the light when leaving the room, and underestimate the effect of larger-scale measures that cause cost or possible discomfort, such as installing energy-saver lightbulbs or reducing the heating.

Finally, there might be problems where the relevant information is not spread out thinly, like in the last example, but concentrated in a few hands (or rather, heads). Here the problem is not logically or computationally complex, the difficulty of finding a solution stems from the fact that these facts may be difficult to access.

Thus, while crowds might be accurate in predicting the outcome of the next election, they are unlikely to be good at predicting whether a certain politician will resign tomorrow for personal reasons or not. This decision (assuming the “personal reasons” are really such and not actually political reasons) is based solely on information in the politician’s head.

Similarly, whether a certain organization of extreme political views is plotting a terrorist attack, or whether they are relatively harmless people with fringe viewpoints, this information is concentrated within the group itself. Clearly the way to deal with this problem is to try and get the information from that source (through bribes, infiltration or electronic surveillance). Trying to model this information problem as a multidimensional information problem will mislead us into thinking that diversity beats ability in this case as well.

Thus, for problems that turn on some piece of knowledge that is costly or otherwise difficult to obtain, but in general simple, diverse deliberative groups have no distinct advantage, and in fact probably a disadvantage, over certain especially qualified individuals.

Again, the issue is that we do not have good reason *either way* to believe or disbelieve that the conditions of the DTA obtain across a wider range of issues. They may

well hold with respect to many political issues that in fact do resemble multi-dimensional estimation problems, and in practical terms it may well be beneficial to increase cognitive diversity in political decision-making (I will actually return to this idea in the next chapter, where I argue that diversity-mechanisms ought to be embedded within an experimental control model of politics). But the standard of political legitimacy is more demanding than that. Therefore: like with aggregative models of democracy, attributing first-order reliability to a deliberative model of politics violates the robustness requirement.

This is a significant point: the implication of the pragmatic standard of political legitimacy is that *no single first-order model of democracy is uniquely justified to be implemented across the board, in political decision-making*. The upshot of this discussion is that we cannot trust any *particular* first-order democratic mechanism to find the right solutions by itself and robustly across the board. Since the uncertainty about the best outcomes is *reproduced* at the level of the uncertainty of the *assumptions* of those models, an adequate model of political organization must be able to employ different institutional mechanisms in an *experimental* manner itself.

#### 4. Experimental *Democratic* Legitimacy

Experimental policy-making therefore may require some institutional experimentation as well, to be successful. The discussion above implies that no single political decision-making mechanisms can robustly be said to be any better than any other at problem-solving. We also may want to introduce markets, commissions, and other forms of independent institutions. This rules out first-order justifications of democracy (especially



those based on “wisdom of the crowds” arguments)<sup>298</sup> more generally. In the end this is not surprising: first-order models of epistemic reliability, even if their reliability is seen as problem-solving capacity (such as in Page’s DTA model), only function within a range of assumptions about the nature of political problems and about people’s capacities and motivations. While the assumptions of these models may be less narrow than for other instrumental justifications of democracy, I have shown that they may not be robust enough for the principle of *pragmatic robust instrumentalism*. One may quite *reasonably*—as I hope I have shown—doubt that a deliberative process such as the one Page and Landemore propose would be optimally reliable at problem-solving across the board.

Now this seems to undermine the project of an instrumental justification of democracy from the start: it seems that no procedure can be instrumentally justified. However, in chapter 3 I have already discussed that an experimental model of *policy-making* may be robustly justified. The problem with the first-order instrumental models of democratic decision-making is that they focus on the wrong *function* of democratic procedures. We cannot robustly make the case that particular democratic procedures are likely to be good at problem-solving. The key move instead is to see the functions of democracy not primarily as “making good decisions,” but as an institutional frame within which different decision-mechanisms—suitable as they may be—are embedded. The function of democratic institutions is not to make all decisions, but to *exercise control over set of institutions tasked with first-order decision-making*. John Stuart Mill put this distinction very eloquently:

---

<sup>298</sup> for instance Landemore, *Democratic Reason*; Dietrich and Spiekermann, “Epistemic Democracy with Defensible Premises.”

“There is a radical distinction between controlling the business of government, and actually doing it. [...] *Some things cannot be done except by bodies; other things cannot be well done by them.* It is one question, therefore, what a popular assembly should control, another what it should itself do. It should, as we have already seen, control all the operations of government. But in order to determine, through what channel this general control may most expediently be exercised, and what portion of the business of government the representative assembly should hold in its own hands, it is necessary to consider what kinds of business a numerous body is competent to perform properly. That alone which it can do well, it ought to take personally upon itself. With regard to the rest, its proper province is not to do it, but to take means for having it well done by others.”<sup>299</sup>

What I call the “Control Model” of democracy is based on this observation. The problem is that, contrary to Mill, it is difficult, and indeed *ex ante* impossible, to determine “what kinds of business a numerous body is competent to perform properly.” As discussed at length above, the fact of *ex ante* uncertainty about this necessitates an experimental approach to policy. The task is now to construct an argument for democracy that is *robust* (i.e. acceptable from reasonable points of view) as well as *optimal* (i.e. it picks out some functionality of democracy that is essential for the experimental model of problem-solving).

I will make this case in two parts. This section 4 of this chapter is concerned with outlining the connection between the functional requirements of a successful experimental strategy of policy-making and the mechanisms of democracy. The argument is that democracy can *robustly* be expected to optimally fulfill those functions within a political system. The next section 5 then connects these functions with the Control Model of democracy—hence a representative form of democracy organized

---

<sup>299</sup> John Stuart Mill, “Considerations on Representative Government,” in *Utilitarianism; On Liberty; Considerations on Representative Government; Remarks on Bentham's Philosophy*, ed. Geraint Williams (London: Dent, 1993) 248. Emphasis mine.

along the lines of the Control Model can robustly realize the ideal of experimental policy-making, and hence, can be justified on the normative grounds of the PRI principle.<sup>300</sup>

*The functional requirements of experimentation*

So what we are looking for is an institutional decision-making mechanism (1) that can be expected to identify political problems and to make functional progress on them, given that there is fundamental uncertainty about what will work, (2) which is also unique to democratic systems of rule, and (3) which will deliver pragmatically reliable outcomes for a reasonable range of possible scenarios.

We have already seen which *strategy* of dealing with epistemic problems is likely to find progressive solutions in the pragmatic sense under conditions of uncertainty. C. S. Peirce calls this the “scientific method,” for Dewey it was the “experimental method” or the “method of inquiry” and in the context of politics we might call this an *adaptive* strategy. The example of the good and bad baker back in chapter 3 already suggested the functional requirements of an adaptive strategy: (1) the identification of problems, (the “identification” function) (2) experimentation, innovation, or creativity—the creation of variation in possible solutions (the “variation” function), and (3) a feedback mechanism such that the success or failure of the possible solutions has an appropriate causal effect back on the source of variation, thus forcing adaptation (the “feedback” function).

---

<sup>300</sup> The next chapter is then concerned with going into more detail how the experimental functions can be realized through concrete institutional mechanisms.

So, if we grant the pragmatic understanding of epistemic adequacy, a decision mechanism would be justified if it fulfils these three functional requirements: identification, diversity and feedback. Can we be sure that democratic mechanisms uniquely do so?

Now this does not seem immediately obvious. Identifying where there are problems within the social sphere could be done through public opinion research; a diversity of possible solutions—as already mentioned—means that we have to consider rather non-obvious ideas and views; and finally, one may think that the relevant feedback that tells us whether a solution has really resolved the problem that sparked it could be determined by some unelected authority just as well as an elected one.

This may in fact be true. However, we have to have a reason to trust that our authoritative regime will *actually* do so. A dictator who is highly attuned to the demands of citizens (and who has a functioning polling agency), or a committee of experts might be very responsive indeed to the feedback from the population. He might even be self-reflexive when it comes to evaluating success criteria. This may be so. But could we make a robust case for this? Among other things, a robust justification requires that the problem-solving capacity should not depend on the goodwill of the people in charge. If I have to trust in an authoritative regime's capacity, that capacity should not be predicated on some particular motivations on the part of the rulers—especially not motivations that I could reasonably doubt they actually have.<sup>301</sup>

---

<sup>301</sup> The equivalent idea, in connection with freedom rather than problem-solving, has been discussed over a number of works by Philip Pettit. See for example Pettit, *On the People's Terms*.

In dictatorships, under rule of experts, and in similar systems, whether identification, variation, and feedback mechanisms work depends on the goodwill and motivation of the rulers. If they are benevolent, or if they are pragmatists, then such a system might be just as adaptive. The crucial advantage of a democratic system is that it works (potentially) even if no-one is especially benevolent, or shows especially adaptive behavior themselves. As mentioned already, democracy is adaptive *at the system level*, and therefore does not depend on political actors or voters having any particular motivation. System-level adaptation can work with angels as well as with regular folk.

The indirect link between the creation of variation in proposed solutions and democratic mechanisms is considered in more detail the next chapter. In this section therefore I focus only the identification and feedback functions.

The identification of problems in a political context has two aspects: first, of course, decision-makers have to be informed about what those problems are. In the previous chapter I have identified a “political problem” in the Deweyan sense as a conflict between “ends-in-view,” that is, a clash between different normative commitments as it is played out in a concrete experiential context, and which requires a collective solution.<sup>302</sup> Thus, a functioning problem-solving must have a pathway such that those clashes can be brought on the political agenda, through, for instance, petitioning the legislature by individuals or non-government organizations, or through the executive branch itself.

Now, there is nothing inherent about democracy that uniquely makes this possible. Authoritarian systems may also include a practice of petitioning, and they may

---

<sup>302</sup> That is, it is not a more general clash of “ideologies” in the abstract. Note that this also is not necessarily a clash between different individuals, or different groups. A “problem” can also occur between an individual and the environment (as it is experienced).

have a very effective system of getting to know the problems of their population.<sup>303</sup> However, the identification function has another dimension: the political system must be expected to reliably *address* every problem that is brought on the agenda: the decision-makers must in good faith attempt to resolve the problem. In the context of the experimental model, as it were, they should be expected to move on to the next step and gather information, in order to create a working hypothesis of how to solve the problem. Now in an authoritarian system, the petitions and other information about political problems sent through the polls have no “teeth”—they do not (necessarily) give any incentives to the ruler to address the problems. In democratic systems, where rulers depend on ex post evaluation, there is a *pro tanto* incentive for deciders to address the problems that are brought onto the agenda.

Elected decision-making bodies have a further advantage in this regard: because of the fluctuating membership, and the open entry conditions, new issues can come onto the agenda even though none of the established decision-makers are liable to change their minds: just simply because they can be replaced with someone else. The remarkable success of single issue parties in Europe over the last decades —Green Parties and Civil-Liberties-Oriented Parties on the left, and Anti-Tax and Anti-European parties on the right—suggests that this agenda-setting function through replacement of established positions is functioning to varying degrees. Of course, ideally, this mechanism of pushing issues onto the agenda may not only work by electing new parties to the legislature, but also by changing the incentives of established decision-makers such that they are more

---

<sup>303</sup> Although a reasonable case can probably be made that authoritarian regimes are systematically bad at knowing the problems of the population, because they induce a fear of complaining (and therefore a “strategic misrepresentation” of beliefs on part of the population).

likely to change their minds. In a two-party context such as the United States, where it is near impossible for a third party to get elected, this same mechanism can function with respect to platforms within the two great parties.

Furthermore, the *publicity* inherent in democratic decision-making forces decision-makers, once an issue is brought onto the agenda, to at least rhetorically *deal* with it. Note that on this understanding we do not have to expect the democratic system itself, or individual decision-makers, to be especially good at *solving the problem*—i.e. we do not claim that democracy is any better at resolving those problems; however, democratic systems can *robustly* be expected to incentivize decision-makers to actually bring problems onto the agenda and address them—either for fear of losing their office, or just at the system level by actually *being* replaced.

The comparison to a non-democratic alternative is illuminating: an enlightened ruler or a technocratic elite may of course also exhibit great concern for their countrymen's well-being. They may even be better at getting things onto the agenda than the messy democratic system that works by cohort replacement and through the incentives influencing people's motivations. However, we only have a reason to believe this if we trust in the ruler's wisdom and motivation: and crucially, it may be *very* reasonable not to do that.

Now of course these mechanisms can work more or less well within a democratic system: politicians are likely to attempt to play down problems that are put onto the agenda, to hide them, or to shift blame to others. Electoral and party systems may make it unnecessarily difficult for new parties or issue groups to emerge and be successful, and

thus give an extreme advantage to established parties (thus preventing the identification of problems). Institutions of the electoral system (e.g. “safe seats”) may even prevent minority opinions and their issues from ever getting on the agenda.

Thus, the argument should not be taken to imply that *all* democratic systems actually fulfill these functions. Rather, understanding the role elections play in this way appropriately directs our attention to their incentive-setting function. This can help us precisely to develop a critical perspective on these particular democratic manifestations. From the perspective of the identification function we can see exactly what is wrong with electoral institutions if they prevent new issues from emerging—even though from a procedural viewpoint they may look completely fair.

In a similar way, we can see the distinct advantage of democracy with respect to the feedback mechanism as well: it has a systemic feedback function built in. The effects of policy have a direct effect looping back to the policy-makers: those making bad policies will lose their jobs, and those who solve problems keep theirs. If (and this might be a big if) citizens base their vote honestly on retrospective evaluation, elections have two functions: they will eliminate policies that have not gotten rid of problems, and they will set proper incentives for parties or candidates to try to solve problems.<sup>304</sup> As such, democratic systems are (potentially) adaptive *at the system-level*.

This is a central point: The system will adapt, much like the system of the biosphere through evolution by natural selection, even if none of the elements of the

---

<sup>304</sup> Of course, there are complications here: what happens when a candidate does not stand for re-election? How is she affected by a looming election? One way this may operate would be that in so far as she is interested in the continued success of her *party* or political movement, she may be take the expected feedback into account.



system do so. None of the politicians have to be especially pragmatic, wise, or even benevolent. If we accept that, we also shift our normative focus: if the main epistemic work is done by the *system-level adaptive nature of democracy*, we should focus less on improving the aggregation of popular judgments or the widest-possible deliberative contestation of different policy proposals. Instead what matters most crucially is the capacity of citizens to give *meaningful* feedback about the effects of policy, feedback that provides *information* about the effects of policy as well as setting the proper incentives that enable the elimination of non-progressive policy: “We should not try to build a better world, we should make better feedback mechanisms.”<sup>305</sup>

As it were, on this theory it does not matter that much *who* makes the political decisions, *as long as there is an effective functional feedback mechanism*. After all: we do not *know* who is more likely to make the right decision before we have tried it. The epistemic role of elections of representatives, for instance, would be a different one: instead of selecting the right representatives; elections make sure they attempt to keep our interests at heart and ensure that the system can correct itself when something has not worked out.

However, this immediately gives rise to another issue: does such a justification pass my own stringent robustness standard as expressed in PRI? After all, is democracy really better than other forms of political organization at realizing this experimental policy-making strategy? For one thing: if scientific inquiry is the hallmark of adaptivity, does this not suggest that rule by experts (perhaps professional “policy engineers”) would be better than democratic rule at *controlling* this process? Perhaps we need to only find

---

<sup>305</sup> Owen Barder, quoted in Tim Harford, *Adapt: Why Success Always Starts with Failure* (London: Little, Brown, 2011), 142. The statement refers to development policy rather than politics as a whole, but I find it an apt summary.

people who have internalized the experimental method of inquiry and let them rule (e.g. “technocracy”)?

Another key alternative is also what we may call the *Hayekian* alternative. If experimentation is what we want, perhaps we should leave as many decisions as we can up for free market processes to work out and sharply restrict the scope of democratic authority? Markets are of course quintessentially adaptive social mechanisms. This is in fact often taken to be their key advantage. If there is no demand for something, if it is produced inefficiently, or if it superseded by a newer product, it will just disappear from the market; a market system eliminates failures faster and (arguably) more efficiently than a democratic system.

However, there is a fundamental difference between science, the economy, and the political sphere. As discussed in chapter 4, since we are dealing with factual as well as normative questions, what is a problem and what constitutes a progressive solution depends on the problem at hand, and the background values. Unlike in scientific inquiry, there is no “natural” criterion (such as explanatory capacity) by which we might distinguish success from failure, and unlike in biological evolution there is no automatic feedback mechanism from the effects of a policy back to the decision mechanism.<sup>306</sup> Markets, similarly, have only one criterion of success: effective demand. Something is good if it meets the immediate preferences of consumers (and if they can afford it), and not otherwise.

---

<sup>306</sup> In biological evolution the “natural” criterion of success is fitness, and the feedback operates by eliminating the unsuccessful before they can procreate.

Technocracy and decentralized market systems are not self-reflexive enough to function as fundamental control systems: if we are not sure what constitutes success, it must be possible to subject the success criterion itself to critical evaluation.<sup>307</sup> Democratic feedback does not use a pre-defined criterion of success: there is no prescription that citizens should employ a specific criterion of success in their *ex post facto* judgment, and indeed no possibility of enforcing any such prescription. If the ends of politics are themselves uncertain, i.e. if we lack “automatic” selection criteria, they must be subject to experimentation as well. Democracy can be appropriately self-reflexive in this way. The only criterion of success is whether individual citizens consider it a success.

Let me note again, though, that markets and scientific committees may well be the right way to address *specific* problem-contexts. But whether they are—whether for example, market mechanisms would resolve problems in higher education<sup>308</sup>—must itself be subject to experimental testing; not least because the *standards of success* for resolving those problems are unclear. As such, markets and scientific committees (as well as other institutional mechanisms) must be embedded within a democratic incentive-setting structure.<sup>309</sup>

---

<sup>307</sup> This is the fundamental argument of Knight and Johnson's version for a second-order understanding of democracy. See Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*.

<sup>308</sup> As for instance the UK government-commissioned Browne Report asserts, see Lord Browne of Madingley, *Securing a Sustainable future for higher education: an Independent Review of higher education funding & Student finance*, 2010.

<sup>309</sup> Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*.

*The Voters*

Now this brings up another question: the argument relies on the assumption that an electoral system that is designed appropriately can robustly force policy-makers to experiment and adapt (whether they want it or not). But this shifts a considerable burden on the voters—who have to actually exercise this control function. Now, can we really expect voters to be up to this task? Is it really *reasonable* to assume that voters have the requisite level of competence?

Relatedly, a number of authors have recently argued that we cannot generally assume that all voters have a satisfactory level of competence necessary to exercise decision-making authority.<sup>310</sup> Regardless of whether those writers are actually correct about this point—it should make us think about whether we can *robustly* assume the requisite voter competence. If we consider their worries not entirely unreasonable, this assumption seems to be a problem.

There is of course some empirical evidence of voters' frequently holding irrational, incorrect, or inconsistent views. In surveys the low levels of political information on the part of the voters comes up time and time again. The divergence between popular views and the views of experts is also well-attested with respect to a number of issue fields: the economy<sup>311</sup> or the global climate.<sup>312</sup> In all of those areas, the significant divergence

---

<sup>310</sup> Caplan, *The myth of the rational voter : why democracies choose bad policies*; Ilya Somin, "Voter Ignorance and the Democratic Ideal," *Critical Review* 12.4 (1998): 413–458; Jason Brennan, "Polluting the Polls: When Citizens Should Not Vote," *Australasian Journal of Philosophy* 87.4 (2009): 535–549; Brennan, "The Right to a Competent Electorate;" however see also Gerry Mackie, "Rational Ignorance and Beyond," in *Collective Wisdom: Principles and Mechanisms* (Cambridge: Cambridge University Press, 2012).

<sup>311</sup> Caplan, *The myth of the rational voter : why democracies choose bad policies*.

suggests at least that the assumption of sufficient voter competence is not as straightforward as it may look.

The empirical evidence is of course open to reasonable debate. On a theoretical level, however, we also may have reason not to expect much of voters' competence because of the possibility of *rational ignorance*. This idea, pioneered by Schumpeter and Downs,<sup>313</sup> basically states that we have no reason to expect people to become competent in political matters, since their vote carries so little impact. It would be irrational for them to expend time and resources on gathering information when their vote will not make a difference within the pool of all aggregate votes, whether it is informed or not. Whether this theoretical issue really arises is an open question.<sup>314</sup> Nevertheless, this possibility ought to give us pause when considering the *robustness* of the experimental justification.

Now, this may indeed be a significant objection to a model of democracy in which voter competence is *directly* causally linked to the eventual decisions. It may indeed be problematic, for example, that voters should be so much at odds with experts with respect to climate change caused by humans, if policy would be determined directly by those voters. I may have good reason *not* to trust the democratic system's problem-solving ability under those circumstances.

---

<sup>312</sup> John Cook et al., "Quantifying the consensus on anthropogenic global warming in the scientific literature," *Environmental Research Letters* 8.2 (2013); Leah Christian, *Continuing Partisan Divide in Views of Global Warming* (Washington, DC: Pew Research Center for the People and the Press, April 2, 2013).

<sup>313</sup> Joseph A Schumpeter, *Capitalism, Socialism and Democracy* (London: Routledge, 2012); Anthony Downs, *An Economic Theory of Democracy* (New York: Harper, 1957).

<sup>314</sup> See for example Mackie, "Rational Ignorance and Beyond;" Russell Hardin, *How Do You Know? : the Economics of Ordinary Knowledge* (Princeton: Princeton University Press, 2009); Richard Tuck, *Free Riding* (Cambridge, MA: Harvard University Press, 2009).

However, this means that the worries about competence mainly apply to *first-order* models of democratic decision-making. The control conception may be much less vulnerable to this form of objection. In the defense of a somewhat different “organic” model of democracy, J.A. Hobson puts this point well:

“The real answer to the claim of lawyers, doctors and the educated classes generally to have more political power because they are better able to use it, is to deny the relevance of their education and ability. [...] The superior knowledge of general politics which lawyers and teachers and other educated classes claim will be no reason for giving them an extra vote; for this general knowledge, if it exist at all, will not be wanted [...]”<sup>315</sup>

Crucially, the experimental model does not require voters to have especially high reliability when it comes to *finding the right solutions*. Indeed, given the uncertainty and complexity involved in policy-making (as has been discussed in detail in the previous chapters), we may just as well doubt whether experts really know better which policy to pursue.<sup>316</sup> Hence, in the experimental model these different levels of expertise on the part of the population truly do not matter. Incidentally, therefore, this also reduces the appeal of plural-voting regimes that give the educated more votes on the basis of their superior competence.<sup>317</sup>

That is, the cognitive task required on the part of the electorate in an experimental strategy is much less demanding. With respect to the identification function, the requirement is that citizens are able to express their problems (clashes in concrete experience) and perhaps to make the connection between their demands and the available political options. The feedback function of democratic institutions requires only

---

<sup>315</sup> J A Hobson, “The Re-Statement of Democracy,” *Contemporary Review* 81 (1902): 271.

<sup>316</sup> For instance, consider the vast disagreement between economics experts on many of the problems today.

<sup>317</sup> The most famous version of this is Mill, “Considerations on Representative Government.”

*ex post* reliability: the ability to distinguish when a solution does in fact work and when it does not—contrast this with the cognitive task of reliably judging whether a solution *will* work. If there is fundamental uncertainty about what is going to work, it does not follow that there is a similar level of uncertainty about whether a given strategy *does* work once we can observe its practical consequences.

We also do not have to judge whether a given policy is an *optimal* solution—only if it has resolved whatever problem prompted it. To illustrate the difference: It is fairly easy to judge whether unemployment has been lowered (whether this can be traced back to a particular policy),<sup>318</sup> while it is fiendishly difficult to judge which policy *will* reduce it, let alone which policy *will reduce it the most*. As Dewey put it, “The man who wears the shoe knows best that it pinches and where it pinches, even if the expert shoemaker is the best judge of how the trouble is to be remedied.”<sup>319</sup> In another instance of the apparently very popular shoe-making analogy J.A. Hobson also responds to the competence objection:

“If it is absurd to suppose that all classes and all individuals are equally wise and good, and therefore equally qualified to contribute by vote and voice to wise and good government, it is not absurd to suppose that every class and every individual knows more about the facts of his own situation than any other class and individual, and can say where the shoe pinches.”<sup>320</sup>

But is *that* really a reasonable assumption? The problem, as noted above, with the diverse first-order models of democratic decision-making was that we cannot assume that majorities generally will come up with better *solutions*—perhaps the same is true with

---

<sup>318</sup> Of course it is often difficult to disentangle possible causes of a reduction of unemployment; nevertheless it is probably comparatively easier than the corresponding predictive task. Furthermore, this could be prevented by testing the policy in a randomized controlled trial before rolling it out.

<sup>319</sup> Dewey, *The Public and its Problems* 207.

<sup>320</sup> Hobson, “The Re-Statement of Democracy” 270.

respect to *evaluating outcomes*? Consider the issue that policies sometimes have very long-term effects. Imagine a politician A who has observed that his city employs too few firemen in case of a natural disaster. He takes some public money and hires more firemen. Most disaster control experts agree that this gives adequate protection. However, nothing happens for several years. The people cannot see the payoff, come to think that A has wasted money and correct the situation—they vote him out. The new incumbent B cuts back on the fire department’s funds. Only the following year there is great tornado damage which could have been contained if there had been more firemen. Now A is vindicated, but it’s too late, the damage is done. Does this not suggest that the majority might well be comparatively bad at giving feedback—especially compared to experts?<sup>321</sup>

One answer to this challenge may be that we have to evaluate the voters’ capacity to evaluate outcomes not *in itself*, but relative to the institutional alternative: Hence, this example obscures two important aspects. First, political decisions are always contextual, involve *tradeoffs* and have knock-on effects. Of course, if there was unlimited money, more firemen would always be better. But we should assume that the money A spent on the fire department was missing elsewhere. Charitably, we can probably assume that B spent the money on some other essential service. The disaster control experts will be able to tell us how many firemen are necessary to contain the fallout from natural disaster, but they cannot tell us how important that is relative to other values. Their expertise should, of course, be taken into account when deciding on this tradeoff, and it should have informed the initial decision by A, but it cannot decide the matter by itself. Second, while the

---

<sup>321</sup> This and similar problems with the issue of institutional progress are brought up in Jon Elster, *Solomonic Judgements: Studies in the Limitation of Rationality* (Cambridge: Cambridge University Press, 1989) 181-201.



citizens may have made a mistake in the example, the question is whether we can expect an expert feedback mechanism to be more progressive across the board. The problem is of course that it is impossible to find these “universal” experts whose feedback will be most likely to recognize progress across the board—and any such assumption will fall afoul of the robustness requirement.

I will accept, however, that if we reject the minimal assumption that people can somewhat reliably judge how their life is affected by a policy, this model would not function anymore. However, seemingly irrational responses to policy choices on the part of the citizenry in many cases may be a function of the obscurity and complexity of rules and policy as they are written rather than people’s cognitive capacities. But if those cognitive capacities should really be so low, this would question the whole idea of democratic, that is, *responsive* politics in the first place. If individuals lack the cognitive capacity to reliably evaluate how outcomes affect *them*, then this is a problem for any common-good oriented political arrangement.

Another issue, however, is that people may have substantially different opinions about the ends of politics, and therefore also what constitutes a progressive solution to a political problem. What if some people prefer more protection in the case of disasters, while others prefer to spend the money on something else? Does that not defeat the argument? Again, recall that the pragmatic criterion is more robust: identifying progress does not require that citizens will reach a consistent full ranking of all possibilities; a partial ranking can be enough—and a partial collective ranking can be reached for a much wider profile of individual judgments than a full one that identified a single

optimum.<sup>322</sup> Thus, if there is widespread disagreement, while there is probably not an agreement on all progressive steps, there might still be large overlap on many progressive transitions. Fundamental disagreement makes finding the optimum very difficult, but we are not looking for the optimum: we are looking for a solution; a democratic system is likely to identify at least some areas of progress under a wide range of profiles.

## 5. The Control Model of Democracy

Now that we have a clearer perspective on the functional requirements of the experimental model of democracy, and therefore also of the functional role the democratic elements of the system should play, we can see more clearly what kind of an institutional form of democracy it would point toward. The first aspect to note is that this model requires a *separation* of the first-order decision-making function (the *generation* of diverse possible solutions) and the ex post evaluation function (the *testing* of those policies).<sup>323</sup>

We can see that this already calls for a *representative* democratic system. In particular, it calls for a representative system not as a second-best approximation to a direct-democratic first-order model of decision-making in its own right. The representatives are there to manage different experimental institutional attempts at

---

<sup>322</sup> See also the discussion in chapter 3 above. For these reasons Amartya Sen for instance prioritizes the realization of progress via incomplete rankings over finding ideal solutions. If I cannot find the optimum, I can still make all the advances I can. See Amartya Sen, *The Idea of Justice*. The same idea is applied to justificatory liberalism by Gerald F. Gaus, *The Order of Public Reason : A Theory of Freedom and Morality in a Diverse and Bounded World*.

<sup>323</sup> For this terminology of "generate" and "test" functions that have to be separated see also Pettit, *On the People's Terms* 203.

problem-solving. As such their role is that of a *deputy*, charged with problem-solving and subject to collective evaluation on the parts of the electorate. In contrast, their primary role is not to *descriptively* represent the population in any particular way (except in so far as this is necessary for the identification/agenda-setting function).

We can see the role of representatives, in this model, is the assembling of what Chalmers calls a *decision network* for each political problem they encounter, which may bring together experts, stakeholders, elected representatives or whoever may be useful to contribute to a problem-solution. The key question is how to *incentivize* representatives in such a way that they are forced to attempt, in good faith, to address these problems. In other words, how do we force representatives to take the problems of the citizenry seriously, and associate the citizens' interest in problem-solving with their own. This is a typical principal-agent problem, or a problem of the alignment of interests.

For James Mill, for example, this is in fact the key issue of representative government: "If things were so arranged, that, in his capacity of Representative, it would be impossible for him to do himself so much good by misgovernment, as he would do himself harm in his capacity of member of the community, the object would be accomplished."<sup>324</sup>

The role of elections is therefore to *control* the experimental first-order decision-making machinery and to force *adaptation* through the identification and feedback

---

<sup>324</sup> James Mill and Thomas Babington Macaulay, *Utilitarian Logic and Politics: James Mill's "Essay on Government," Macaulay's Critique, and the Ensuing Debate*, ed. Jack Lively and John Rees (Oxford: Clarendon Press, 1978) 75.

functions.<sup>325</sup> This second-order model of democracy, which emphasizes the *oversight* or *control* dimension of democracy is the centerpiece of Jack Knight and James Johnson's pragmatist version of democracy, and is also at the heart of Philip Pettit's republican conception of democracy.<sup>326</sup>

Applied to the experimental model of policy-making, the main epistemic function of democracy is not to make reliable first-order decisions, but to manage the different institutional arrangements that make those first-order decisions. As mentioned: this management involves both incentive-setting to keep those involved in first-order decisions "honest," but also the reflexive assessment of the success of those first-order decision mechanisms: "The priority we accord to democracy reflects its usefulness in approaching the crucial second-order tasks involved in the ongoing process of selecting, implementing, and maintaining effective institutional arrangements."<sup>327</sup>

Knight and Johnson locate the advantage democracy enjoys with respect to this second-order task in three institutional mechanisms: first, in the fact that periodic voting provides incentives and feedback that enables the adaptation of first-order decision procedures; second, in the fact that the openness and publicity of democratic argument forces initiatives to face scrutiny from a varied public; and finally, in the fact that democracy is an inherently self-reflexive, and therefore adaptive system that subjects decisions to constant revision and re-examination.

---

<sup>325</sup> For the democratic function of "control" see also Kitcher, "Public Knowledge and the Difficulties of Democracy" 1212.

<sup>326</sup> Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*; Philip Pettit, "Depoliticizing Democracy," *Ratio Juris* 17.1 (2004): 52-65; Pettit, *On the People's Terms*.

<sup>327</sup> Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism* 260.

In more concrete institutional terms, what this means is that we have to evaluate our democratic institutions with respect to three questions (and relative to their alternatives): (1) Does the openness and transparency of the system reliably force representatives to identify areas of popular dissatisfaction (the presence of “problems”) and force them to publicly address them? (2) Does the system reliably encourage—or even *force*—the creation of a diversity of solution proposals? And last, but definitely not least, (3) do elections *reliably* express the people’s evaluation of the success of representatives’ problem-solving attempts? These are the core criteria of the experimental model of democracy.

These aspects are examined in the next chapter, which suggests that democratic mechanisms *can* indeed be designed in a way to satisfy these requirements. This, then, is the normative model of democracy that satisfies Pragmatic Robust Instrumentalism: “Experimental Democracy.” Recall that the dilemma that motivated the dissertation was that we are deeply concerned with outcomes, while we do not know how to *robustly* assess the reliability of different first-order decision mechanisms. The control model of democracy can be reasonably expected to deliver problem-solving—not on each and every decision it produces, but over time and in an adaptive fashion. Furthermore, it does so *without* relying on controversial (or potentially controversial assumptions). A democratic system based on the control model can therefore be *robustly justified on an outcome basis*. Reasonable people who deeply disagree, still have a reason to endorse this political arrangement.

This justification differs therefore both from arguments that the premises of justificatory liberalism and reasonable pluralism imply a commitment to a purely

procedural understanding of democracy on the one hand.<sup>328</sup> On the other hand, however, this argument is also markedly different from other versions of the control conception that ground its legitimacy in potentially controversial normative theories, like utilitarianism<sup>329</sup>, republican freedom as non-domination,<sup>330</sup> or pragmatic epistemology.<sup>331</sup> I am of course sympathetic to the institutional conclusions of those arguments, however, I hope I have given a more stable and robust basis for such a conception, based both on the value of political outcomes *and* on the fact of deep disagreement and uncertainty in the realm of politics.

The next chapter addresses in more concrete terms how the functional requirements of *experimental democracy* can be realized within a system of representative democracy that is based on the second-order control conception.

---

<sup>328</sup> Waldron, *Law and Disagreement*; Valentini, “Justice, Disagreement and Democracy;” Urbinati and Saffon, “Procedural Democracy, the Bulwark of Equal Liberty.”

<sup>329</sup> Mill, “Considerations on Representative Government.”

<sup>330</sup> Pettit, *On the People's Terms*.

<sup>331</sup> Knight and Johnson, *The Priority of Democracy : Political Consequences of Pragmatism*.

---

## The Institutions of Experimental Democracy

### 1. Introduction

The last chapter has tried to establish that a representative, “control” model of democratic decision-making is in general appropriate to realize the ideal of experimental democracy. This chapter considers a number of concrete questions about how a more experimental model of policy-making could be realized within the context of the control conception of democracy.

Remember the functional requirements of experimental democracy, as defined in chapters 3 and 4: the identification of problems, the creation of a variation in possible solutions, and especially the ex-post evaluation and consequent adaptation. As the literature on decision-making under uncertainty as well as the pragmatists tell us, this is an adequate way to address political problem situations under conditions of complexity.<sup>332</sup>

---

<sup>332</sup> Axelrod and Cohen, *Harnessing complexity : organizational implications of a scientific frontier*.

In the last chapter I have argued that in principle, on the basis of the pragmatic justification, a representative “control” conception of democracy is a justified form of rule. This is because we cannot make any robustly justifiable judgments about the first-order capacity of democratic institutions to deliver pragmatic problem-solving. The institutional strategy must be “experimental” in that the policy-making body must be free to pursue experimental strategies with respect to policy problems, both in terms of policy and of institutional design. The role of the electorate in this conception is to exercise oversight, keep the policy-originators on track (and remove them if need be) and through this mechanism force adaptation. As Pettit puts it, the electorate fulfils a double role here: to give a general direction to policy-making in the long and medium term, and in the short term provides feedback and incentivization to keep policy-makers on that direction.

333

This chapter considers a number of institutional mechanisms that would bring together the control conception of democracy and the functions of experimental policy-making. After all, even if we accept the pragmatic theory of democratic legitimacy, how can we be sure that any given political decision structure actually fulfills the criteria for this model? The major question is therefore this: can the experimental model of policy-making be realized within a representative system, and if yes, what institutional features are necessary for that? After all, theoretical considerations do not automatically mean that it is possible to create a stable and robust institutional realization of experimental democracy.

---

<sup>333</sup> Pettit, *On the People's Terms* ch. 4.



For that reason, this chapter is organized as a series of responses to objections like, “the ideal of experimental democracy sounds nice, but there is no way to robustly realize it.” The goal of the present chapter is not to come up with the *definitive* institutional blueprint that would be uniquely desirable under the conditions of robust instrumentalism (that would of course be presumptuous). Rather, in the spirit of this whole project, I acknowledge a significant degree of *uncertainty* about this question (which, of course, calls for a degree of *experimentation*). The answers given in this chapter should therefore be taken as illustrations of the *kinds* of institutional responses to the functional demands of this particular normative theory of democracy may work. As such, they are both there to show the *possibility* of institutionally realizing experimental democracy, and to tentatively establish a critical view from which one may evaluate *existing* democratic structures.<sup>334</sup>

### *The Basic Issues*

Thus, going back to the three elements of the experimental strategy, we need to think about the institutional realization of the following aspects: (1) The *identification* of political problems; which includes the signaling function that there *is* such a problem at a particular location in society, but also a mechanism to put them on the political agenda and encourage/force a good faith attempt at experimentally resolving them. (2) The *creation of variation in potential problem-solving solutions* as well as *putting problems on the agenda*. As the last chapter laid out, the key issue with direct and more strictly majoritarian approaches (mandate-driven representation for instance) was that we could not be sure

---

<sup>334</sup> See also chapter 1, section 3 for a discussion of these goals.

where problem-solutions or problem-reconceptualizations would come from. Input from a variety of sources is necessary. (3) Finally, we need *feedback* from the population that fulfills the double role of incentivization and information transmission. The vote will only do so when we can be confident that the collective judgment expressed in the majority vote really does reflect a reasonably sincere and non-strategic evaluation of their individual problem situation, that the outcome is not an artifact of electioneering or the vagaries of social choice paradoxes, and if the vote really carries meaningful information as to the collective evaluation of the relative performance of different policy-makers. It must be an accountability mechanism that focuses on the actual success or failure (the “performance”) of the policies enacted.

As it were, unless we can reasonably trust that our particular democratic institutions actually do function in this way, they are not justifiable according to the pragmatic model. These criteria therefore give us a perspective to judge existing democracies as well as different mechanisms within them. However, as I shall argue, these conditions are in principle relatively easily fulfilled. By this I mean that they can be fulfilled mostly on the institutional design level alone.<sup>335</sup> In the remainder of this chapter, I will briefly address all three points above. In sections 2 I address the identification function, and consider how we may bring a political procedure to robustly address political problems. In section 3 I consider the question how a variety in problem-solving approaches can be brought to bear on the political process. I argue in line with some recent work on this topic, that we should understand this as the question how efficiently to bring *cognitive diversity* into the political decision process. I argue that we cannot expect

---

<sup>335</sup> They do not, for instance, require great changes in citizens’ beliefs or motivations.

democratic institutions to directly create this diversity, let alone to bring it to bear on the policy process; *cognitive diversity* needs to be actively selected for. In section 4 I consider publicity and reporting requirements on the legislative process that may enable the standard-setting and monitoring necessary to enable a controlled implementation of policy. In that section, finally, I also consider the question of how to make the collective judgment expressed through the vote more meaningful and effective. Here I will rely especially on Michel Balinski and Rida Laraki's idea of replacing the first-preference vote with a system of collective evaluation, or as they call it, *collective judgment*. This simple change of the mode of expressing collective choice has a number of extremely valuable benefits on all the required dimensions.

This chapter considers how to bring *cognitive diversity* into the political process, how to ensure *clarity and publicity* of the policy-making process, and how to make the democratic accountability mechanisms more *effective and more informative*.

This chapter will therefore serve as an illustration of how experimental democracy may be possible; I do not claim that these are the *only* ways in which we might bring this into existence, but it shows that in principle, democracy can be a reasonably justified and legitimate form of political authority—even with the very demanding standards with which this dissertation began: the high standard of the combination of justificatory liberalism and of the requirements of outcome-based theories of political legitimacy; and the fact of pluralism and deep disagreement as well as pervasive uncertainty about the expected consequences of political action.

## 2. The Identification of Problems

The first requirement we should consider is the identification function. Recall that in the control model of democratic politics, the assembly of representatives is free to employ experimental approaches to problem-solving, including experimenting with institutional structures. Thus, depending on the political field at hand, first-order decision-making authority may be delegated (conditionally and reversibly) to committees, external agencies, ad-hoc *decision networks*, or even market forces.

This model of politics depends crucially on the right problem-solving motivation and good faith efforts of the assembly of representatives to resolve the problems of the population. Hence, this whole arena of institutional and policy experimentation is embedded within democratic institutions that give direction to and exercise control over the representatives' attempts at problem-solving.

The skeptic may now object that it is by no means given that, even if they are subject to periodic elections, the representatives would come to address the problems of the population. This is especially pertinent when the problems concern only a minority. This is of course a huge problem with respect to more "permanent" minorities such as special identity or interest groups; but this is in fact a more general issue with any problem that only a minority of the population happen to care about (even if that is not a stable group). Why should those in charge respect wishes of the minority if they do not depend on their electoral backing? In first-past-the-post systems, or generally systems that tend to disproportionately overrepresent majority opinions, this is exacerbated.

In a proportional representation system of course, the legislative minorities depend on the electoral support of their respective minority bases. They cannot afford to ignore the problems of the people they are representing, even if it should only be a few of them. That is already one reason why we might prefer a proportional representation electoral system over one that tends to produce two- or two-and-a-half party systems.

However, from the perspective of the political decision-making body as a whole, the legislative minority factions can of course be themselves ignored in the policy process. If we run parliamentary procedure on a majority rule basis, then while minority viewpoints may have an advocate in the legislature, there is no robust guarantee that they will causally affect the agenda. Hence the problem: while we may trust in experimental democracy's problem-solving capacity when it comes to problems that affect everyone (or at least are obvious to everyone), we may have less reason to trust that they will identify the smaller (in scale, not importance) problems. It is not an option to point out that these will tend to be ignored by all political systems. Robust legitimacy depends at a basic level on the notion that *everyone* has reason to endorse the authority of our political mechanism.

In this section I will therefore talk about two institutional mechanisms and rules that potentially address this issue. The goal is to suggest that it may be possible to address minority this problem within experimental democracy; it is not that these proposals are exclusively the best decision mechanisms. The first is the question of the *entry* of new contenders into the political competition, and the second the possibility of *submajority rules*.

### *New Contenders*

The first way to get a political problem on the agenda is the emergence and representation of a new party in the legislature. Political parties do not necessarily have to be “ideological” in the traditional sense—that they represent a certain socio-political worldview that provides more-or-less coherent answers to a broad spectrum of political issues (e.g. a certain position on the left-right spectrum). The latter half of the 20<sup>th</sup> century has witnessed the rise of single-issue parties in a number of countries. These parties tend not to provide a full spectrum of political positions, but often already in their name give away the focus on one particular problem area. The most successful example of this are Green parties, which over the last decades have gained enormous influence in many European political systems.<sup>336</sup> This is exercised not only directly, through causal influence on the decision process, but also indirectly through the bringing of environmental issues onto the agenda—and thus changing the program of the “ideological” parties to include potential solutions to those problems. Green parties have only occasionally had direct political influence, mainly as minority partners in government coalitions. The emergence of Green parties as serious electoral contenders may therefore be seen as an example of the *identification* function at work, through the dual path of actually replacing existing members, or through changing their political program because of the challenge they represent.

However, this phenomenon goes beyond the incidental success of the Green parties in bringing environmental problems onto the agenda. Some examples of single-

---

<sup>336</sup> Even the anti-proportional UK parliamentary system since 2010 has had one Green member, Caroline Lucas, MP for Brighton & Hove.

issue parties in Europe are: parties concerned with technology and digital rights, especially privacy concerns and intellectual property law (the so-called Pirate parties, which have been moderately successful in Germany and the Northern European countries), anti-Welfare-State-Reform parties (WASG/Left Party in Germany), anti-tax pro-market parties (the so-called Progress parties in Norway and Denmark), anti-European and anti-Euro-currency parties (The UK Independence Party, the AfD in Germany, Syriza in Greece, Freedom Party in the Netherlands), parties representing regional interests as well as some rather unsavory xenophobic parties (Freedom Party in Austria, Front National in France, Vlaams Belang in Belgium, Danish People's Party, the "True Finns").

Whether or not it is a good thing from an objective standpoint that these small single-issue parties are moderately successful is of course an open question. There is, however, no doubt that the entry of these parties can significantly change the political agenda.

The conclusion we may draw is that a well-functioning experimental political system should encourage the entry of new contenders into the political competition. Note that this is not necessarily equivalent to say that we should increase their chances of *being elected*. However, in order to bring their problems on the political agenda, citizens have to have a wide choice of political expression—that is, a wide range of parties for whom to vote. New contenders can also change the conversation.

There are a few possible ways in which the barriers to entry by new parties may be reduced. An obvious candidate is public campaign funding, which means that new

parties can overcome the obstacles of having to compete for funds with well-established players; another measure could be a low threshold for parliamentary representation, such that even smaller single-issue parties can hope for some degree of representation. Finally, we may think about limits in direct or indirect campaign advertising, or even proportional representation rules for public broadcasting. All of these rules initially seem rather peripheral to the democratic idea. From the perspective of experimental democracy, however, these issues are right in the center.

Which measures are in the end likely to be successful is a question that is beyond this project. What we can take away from this brief discussion is however, that a legitimate democratic system should be organized in such a way that minorities can exercise influence over the political agenda through a choice from a wide range of parties.<sup>337</sup> This gives us a critical perspective from which to evaluate existing political institutions as well: does the electoral system/the system of campaign financing/etc. encourage or discourage the entry of new contenders or cement the advantage of established parties, such that people seeking change will have to work through the established parties and coalitions? There is nothing within the conceptual core of “democracy” as such that would help us decide what we should think about these questions. It is not the case that one of those is “more democratic” than the other. However, we can say which of those is going to be more effective in promoting the experimental problem-solving function of political institutions.

---

<sup>337</sup> This may immediately lead to the question of what people do who are concerned with more than one problem. I will return to this issue in section 4 below.



### *Submajority Rules*

So now (ideally) we have some people in the legislature who are interested and invested in some of the problems that concern only a minority. The next problem is how to get the legislature as a whole to address those problems. Parties elected on a particular platform can probably be expected to push for a resolution of the particular issue (at least in order to advance their chances of re-election). They will push for assembling one or more decision networks to address the problem at hand. But can we expect this to go through?

In a strict majoritarian system, we should probably have our doubts. The problem is not only that issues will be voted down by a majority of members in a public vote. An even more acute danger is that problems will not be put on the agenda, or will be “killed in committee”—i.e. rejected without public acknowledgement.

There needs to be a mechanism, therefore, to systematically get what we have called political problems onto the policy agenda. One effective way to do this would be through allowing *submajority rules* within parliamentary procedure. This is a catch-all term for a set of rules that allow a minority (of legislators, or citizens) to initiate political action. A submajority rule does not mean that a policy proposal can be approved by only a minority of representatives. Rather, these rules tend to *force* an issue onto the agenda, that is, a minority decision may force the *public* acknowledgement of an issue, and require an official response; if not a positive one, at least a reason why the problem does not need to be addressed.<sup>338</sup>

---

<sup>338</sup> Vermeule, *Mechanisms of Democracy: Institutional Design Writ Small* ch. 3.

They are, as it were, rules that “force majoritarian accountability.”<sup>339</sup> Within legislatures, these may call for an up-or-down vote on a given policy proposal that can be initiated by a minority. Outside the legislative institutions, submajority rules may allow for the initiation of a referendum, recall or popular veto by a minority of the voting population. Beyond that, minorities may call for official responses or the publicization of data or information through official queries.

Now of course submajority rules do not guarantee that an issue will actually be addressed: after all, anything that is put on the agenda by a minority can subsequently be voted down by a majority. However, by forcing a public response from the majority of the legislature (or majority of the population in the case of a referendum), they make it less likely that the problem can be ignored. Coupled with a publicity requirement therefore, one may hope that the “civilizing force of hypocrisy”<sup>340</sup> would prevent self-serving dismissals of the problem as irrelevant. The disciplining force of “candor” on policy-makers’ behavior may in many cases be more effective than the institutional mechanisms that enforce accountability, especially since elections are relatively indistinct and “blunt” instruments for this task.<sup>341</sup>

There are potential issues with this proposal, of course. We have to specify how large (or small) the minority has to be to put an item on the agenda, and to force a majority response. Stability has to be weighed against the ease with which minority issues

---

<sup>339</sup> *ibid.*, 92.

<sup>340</sup> Elster, *Securities against Misrule*.

<sup>341</sup> For an argument along those lines, see Jeffrey E Green, *The Eyes of the People : Democracy in an Age of Spectatorship* (Oxford ; New York: Oxford University Press, 2010).

can be brought into the political process.<sup>342</sup> Furthermore, submajority rules of course do not *guarantee* that problems are put on the agenda, let alone that they are addressed. An entrenched majority may decide to give spurious reasons for their dismissal of a minority claim. In a particularly antagonistic political system (where political gains can mainly be made at the expense of the opponents) the majority may even benefit from public dismissals of minority claims. This illustrates that none of these mechanisms will guarantee accountability in the sense of the identification function by themselves.

However, we can see that in combination with some complementary institutional mechanisms it is possible in principle to find a mechanism through which problems are at least more or less systematically *put on the agenda*, and there is a mechanism that forces the official acknowledgement of the problem by the legislature as a whole—and nothing more is required in this context from the institutions of democracy at this point.

### 3. Creating Diversity in Problem-Solving

One of the essential requirements of an experimental approach to problem-solving is that a lot of different potential solutions are *generated*, some of which can be tested through the experimental policy-making approach. The key difference between what in chapter 3 was called the “juror” and the “scientist” model of politics was that in the latter, the key task was to come up with appropriate progressive solutions to real-life acute political problems, rather than to come up with a correct verdict.

---

<sup>342</sup> Vermeule, *Mechanisms of Democracy: Institutional Design Writ Small* 101.

A key institutional-design issue with this claim may be that the *control* version of democracy does not seem to encourage the requisite cognitive diversity that is a precondition of the *variation* function. As pointed out in the last sections of chapter 5, the composition of the first-order decision-making institutions within the control model is left open. There is no expectation that elections (understood as feedback mechanisms) would actually bring about cognitive diversity in the set of deputies, and so it seems that we should not expect that cognitive diversity is effectively brought to bear on the decision-process itself.

As it were, we may think that if we want to create variation in solution proposals, we have to employ some form of *first-order* democratic decision-making that actually *embodies* maximal (or sufficient) cognitive diversity within democratic institutions themselves (i.e. in the representative legislature). We can see that this demand may be at odds with the feedback function of elections. If we encourage voters (or push them through institutional design) to increase maximal cognitive diversity among representatives, this may conflict with our desire to get them to evaluate policy performance in their voting choice. For example, the creation of diversity may require that voters choose candidates on *group identity* grounds (i.e. vote for people who are most like themselves); and this, for example, may make it difficult to get rid of unsuccessful legislators—unless there is always a challenger from the same group available and running. The two functions may conflict in the application.

In response, this section defends the idea that *effective* employment of cognitive diversity in political decision-making is not only compatible with the control model of democracy, the control model is actually *better suited* to do this. In order to make this point,

I critically examine the argument that the epistemic benefits of cognitive diversity call for a first-order model of democracy. This claim that the epistemic benefits of cognitive diversity *justify* (democratic) political institutions that create/promote/utilize this cognitive diversity, I will call throughout the *Cognitive Diversity Argument for Democracy*.

It may be important to note that I do not dispute that cognitive diversity has great epistemic benefits. It is in fact a prerequisite for the experimental model of politics. Given the complexity and uncertainty of the social, political and economic problems with which we are faced, this seems a relatively incontrovertible view. It has had strong support at least since Aristotle's *Politics* and the well known discussion in Book III, chapter 11, and it also seems just intuitively plausible to me. Added diversity could markedly improve decision-making in many areas, including the political—and to the extent that it is possible, diversity ought to be encouraged and utilized in the political process.<sup>343</sup>

However, what I do want to argue is that a commitment to the epistemic value of diversity *does not automatically entail* a commitment to a first-order model of democratic politics. As I show in the following argument, we should have our doubts that democracy as a first-order decision mechanism does *actually* encourage diversity, let alone utilize it in an optimal way. Whether or not democracy promotes diversity depends *to a very fundamental level* on more-or-less contingent social circumstances. Furthermore, these doubts are systematic and not merely practical worries about real existing democracies. As I argue, there is no particularly strong reason *in principle* to think that democratic mechanisms *as such* will be particularly efficient at employing diversity in their decision process at the first-order level. The point is not that democracy is *incompatible* with

---

<sup>343</sup> This has of course also already been addressed in the previous chapter.

diversity—not at all—however, ensuring an adequate level of diversity may require *intervention* within democratic procedures, interventions which are not themselves democratic.

Hence, the argument that democracy will make better *first-order* decisions because it utilizes diversity in an efficient and appropriate way, fails. For that reason, as I show, democracy understood as a *second-order* indirect “control” institution can have a role in actively bringing some form of diversity to bear on the decision-making process.

### *The Background*

Recently, there have been some powerful arguments showing the epistemic advantages of certain forms of diversity. First and foremost, as already discussed at length in the previous chapter, there is Scott Page’s *Diversity Trumps Ability* Theorem. Another strong argument in favor of this view can be found in Josiah Ober’s magisterial study of decision procedures of democratic Athens, *Democracy and Knowledge*.<sup>344</sup> Both of these works, it seems to me, endorse a similar enough point that they can be treated as variants of the same general type: (a) that, *ceteris paribus*,<sup>345</sup> the interaction of a diversity of cognitive skills within collective decision-making greatly facilitates collective problem-solving; connected with this are the secondary and implicit points that (b) that increasing a *political* procedure’s

---

<sup>344</sup> Page, *The Difference*; Ober, *Democracy and Knowledge: Innovation and Learning in Classical Athens*.

<sup>345</sup> That is, assuming the added diversity does not introduce secondary problems—e.g. issues of coordination, or conflicting interests. It is an open question whether the *ceteris paribus* condition is a realistic one, but that question must remain unanswered for now.

problem-solving capacity is a good thing, perhaps even a crucially important thing,<sup>346</sup> and that therefore (c) more of that diversity is better than less (within reasonable limits).

As mentioned above, I do not here question any of these three principles or the methods by which these authors come to them. For the purposes of this chapter I shall just assume that they hold, without providing further argument for them.

In order to make the connection between diversity and democracy here, we would have to show that first-order democratic decision procedures optimally (or sufficiently) employ cognitive diversity, and that they do so in the *robust* way required of this type of argument, which has been made in a number of recent works in democratic theory,<sup>347</sup> relies on the claim that democratic procedures can (or tend to) reproduce the mechanisms of cognitive diversity that Page and Ober, among others, postulate. Sometimes this is just assumed;<sup>348</sup> however, this chapter tries to examine precisely this assumption and whether we are justified in making it.

In the following my answer will be a qualified “no.” By this I mean that while first-order democratic institutions *may* produce and utilize cognitive diversity, clearly not all forms of democracy will do so—whether or not democratic procedures as such use diversity appropriately depends on a number of conditions. In particular, there is also an alternative: I suggest that if we want cognitive diversity to have an impact on our policy

---

<sup>346</sup> To be fair, Scott Page does not primarily focus on this latter point. His argument is more general, and while he explicitly endorses more diversity in political decision-making, he does not enter the debate as to the relative value of problem-solving capacity versus other considerations.

<sup>347</sup> For instance in Landemore, “Why the Many Are Smarter than the Few and Why It Matters;” Landemore, “Deliberation, Cognitive Diversity, and Democratic Inclusiveness: an Epistemic Argument for the Random Selection of Representatives;” Landemore, *Democratic Reason*; Colin Farrelly, “Virtue Epistemology and the ‘Epistemic Fitness’ of Democracy,” *Political Studies Review* 10.1 (2012): 7–22.

<sup>348</sup> for instance in Landemore, *Democratic Reason* 7.

process, we should *actively* select for it. We might need to tweak our procedures in such a way as to overrepresent diversity, or to create diversity where it is lacking—and in the process we may have to override democratic precepts. Furthermore, as I point out below, *which kind* of diversity we need is problem-specific and *ex ante uncertain*. Thus, the presence of cognitive diversity depends on active selection and an experimental approach. Therefore, the control model, which allows for the experimental employment of task-specific decision networks on a problem-by-problem basis, under second-order democratic supervision, is actually *better* suited to fulfil the *variation* function by bringing in cognitive diversity.

This section proceeds by addressing three major areas, or questions, that bring out this difference, and illustrate the advantage of the second-order model over the first-order one. The three questions I am looking at are as follows:

- (i) What is the kind of diversity we need and how to ensure we have the right kind? (The “granularity” question)
- (ii) How is the diversity brought to bear on the decision process? (The “mechanism” question)
- (iii) Where and when, in the political process, do we want this kind of diversity? (The “application” question)

### *Diversity of What?*

First we have to be clear *what kind* of diversity we mean when we consider the value of diversity in coming up with a variation in creative solutions. The answer is not



immediately obvious: it could be diversity of age, life experience, formal education, *knowledge*, opinion, ethnicity, interests, socio-economic background, or any number of other indicators. While it is true that these sometimes go together (e.g. that an ethnically diverse group may also be likely to be diverse in a socio-economic sense), we have to know what kind of underlying diversity would be doing the epistemic heavy lifting in terms of political problem-solving.

### *Information Availability*

A first possible reason form of diversity has to do with *information availability* within the decision procedure. Sometimes, good collective decisions may depend on a diversity of *factual knowledge*. Frequently, political problems tend to be difficult and have complex and non-obvious solutions. In such circumstances, there might be benefits in aggregating a diverse set of knowledge within a single process. A group of individuals with not completely overlapping sets of knowledge might be better at finding the right solutions to given problems. Of course we cannot select a group for “knowledge” in the strict sense of “true belief” (unless we already know what is true, which defeats the purpose), since people might hold all sorts of erroneous beliefs in addition to items of knowledge. And that is not what is generally meant by this: what we can do is select for a diversity of *justified opinions*.<sup>349</sup>

---

<sup>349</sup> After all, people cannot really have conflicting *knowledge* about, say, the effects of greenhouse gases on the global climate; but they may well have different justified *opinions* about it, one of which may even turn out to be true (and hence knowledge)

There might be different reasons for that. First, if there is a right answer, but we do not know what it might be, increasing the diversity of opinions might increase the likelihood that the correct (i.e. true) opinion is represented within our decision-making collective as well. Even though in such a case we strictly care only about the *one right piece of information*, if we cannot identify it ex ante, we should probably try to get in as many opinions as we can. We should focus on including *dissenters* in our decision procedure; that is, people who hold beliefs different from the rest of the group.

Second, the right answer might depend on a number of different pieces of information that have to be combined and that are likely to be held by people with different opinions. A decision procedure where many different opinions are represented might be more likely to find and fit together the relevant pieces of information. It might be that nobody fully grasps the problem (and the solution) by themselves. For instance, successfully building a house depends on combining the practical knowledge of the architect, the structural engineer, the bricklayer, carpenter, electrician and many others. What is needed in that case is a *diversity of professional expertise*. Similarly, designing a policy may depend on the aggregate knowledge of a number of different people: for instance, people who know what effects a certain policy will have on different population groups. The difficult question is here, however, which *type* of knowledge diversity is the relevant one. For the construction of houses, on the other hand, the relevant diversity is pretty clear. Formally, these types of problems can be represented as estimating an unknown complex multi-dimensional model, such that it is unlikely that any one individual can

hold a complete model. The group, on the other hand, might combine their respective estimates into something that more closely approximates the true state of the world.<sup>350</sup>

And third, the answer to our political problem might depend on aggregating a lot of latent knowledge that is *widely dispersed*. This is the classic problem of economic planning: a hypothetical Soviet economic planner needs to know which products everyone in society would like to consume, and also the relative efficiencies of the different producers of a given good. Producing a coat that no-one wants to buy, or assigning people terrible at sewing to the coat factory would be the classic mistakes. This problem—which of course is the key motivation for Hayek’s idea of social epistemology<sup>351</sup>—could be generalized to *preferences* as such. Even a benign lawgiver who wants to fulfill the preferences of citizens needs to know what people would like—not only in the economic sense, but also in terms of regulations and decrees—and this knowledge is of course distributed across the heads of all citizens. Short of having everyone’s personal input for every political decision, a diverse decision-making procedure that represents many different opinions about what people might want from government, might mimic this procedure and aggregate enough dispersed knowledge.

If we believe the main issue is this lack of access to the right pieces of knowledge or *information*, then we can see how inviting contrary opinions and people who disagree might turn out eventually beneficial. These ways of thinking about the epistemic benefits of diversity in terms of factual knowledge are relatively typical, as they correspond to everyday experience (like the example of the house construction used above). And it

---

<sup>350</sup> Page, *The Difference* 341.

<sup>351</sup> Friedrich A von Hayek, “The Use of Knowledge in Society,” *The American Economic Review* 35.4 (1945): 519–530.

seems to me that generally speaking, it holds in the political realm as well that good decisions require a diversity of factual knowledge. So does this mean that first-order decisions ought to be democratic?<sup>352</sup>

This conclusion does not actually follow. Firstly, it is not obvious that *physical presence* of the relevant factual diversity at the point of decision-making is really necessary. In theory, a person with a laptop and an internet connection can access all the opinions he needs (and more). He might not know what to look for, and even more crucially, he might not be inclined to consider different opinions; but if the issue is merely the *availability* of the right information, then given current technology for an opinion to be registered and considered within a decision process it is not *necessary* for someone holding that opinion to be physically present within the room. In ancient and early modern times the only way to find out what was going on in a given part of the country, or among a certain group of people, was to send a representative from that region to the capital (preferably with written instructions) but this is strictly no longer necessary when access to information is so instantaneous and virtually costless. One might of course argue that unless there is an advocate physically present, dissenting opinions are liable to be ignored—however, this is a matter of degree. Surely there are opinions so unlikely (think “9/11 was an inside job”) that are liable to be ignored even if an advocate should be present. Whether or not an opinion will be considered within a collective decision process surely depends not only on whether someone is in the room advocating it, but also largely on the importance decision-makers accord it. And this itself depends on the motivations

---

<sup>352</sup> Landemore, *Democratic Reason*.

(and therefore incentives) of those decision-makers. This also ties in with the incentivizing role elections play within the control model.

### *Information Processing*

However, perhaps the issue goes beyond *information availability*. Perhaps the key is rather the *information processing* capacity of the collective. That is, even if all the relevant information is just one Google search away, the key problem is how to distinguish the good from the bad, how to analyze and amalgamate all the different pieces of information and how to craft the right piece of legislation from this huge pile of data. This is the key insight we can take from both Page and Ober's analyses. Both point out that in order to make good decisions, we need the right combination of practical *cognitive capacities*, not only the right knowledge. The diversity we need is therefore a *diversity of problem-solving skills*. As they show, certain types of diversity have very powerful benefits when it comes to information-processing.

Page—parts of whose theory have already been discussed in chapter 5—considers four types of diversity that contribute to this processing capacity: These are: *diverse perspectives*, that is, representations or ways of looking at the of the world; *diverse interpretations*, that is, broadly speaking, ways of categorizing one's representations of the world; *diverse heuristics*, that is, ways of generating simple action-guiding problem-solving rules from one's interpretations; and *diverse predictive models*, that is, ways of inferring cause-and-effect relations from one's representations.<sup>353</sup> As Page demonstrates, when it comes to

---

<sup>353</sup> Page, *The Difference* 7.

adequately assessing multi-dimensional problems, creatively solving those problems, and generating accurate predictions, diversity on these four capacity dimension has enormous potential benefits. Ober shows in a similar way (if implicitly) how the institutions of democratic Athens had the effect of generating such forms of diversity and crucially also the interpersonal ties that enabled their combination within the decision process, but also the diffusion of the new solutions back through society. As Ober argues, in no small way this contributed to making democratic Athens the dominant power in Greece for over a century.<sup>354</sup>

Of course this only makes sense if political problems we encounter are multidimensional and cognitively difficult. More specifically, they have to be *complex*.<sup>355</sup> If problems are easy (any person selected at random could solve them), or if they depend on difficult-to-obtain but essentially non-complex information (e.g. does terrorist group A plan an attack in the next months, and if yes, where?), diversity will not necessarily show its benefits. Similarly, if problems are difficult, that is, if they require a lot of computational capacity, but if this difficulty is a result of a limited set of problem dimensions that are clearly defined, diversity is not necessarily useful. For instance, finding the right solution to the problem of, say, improving school performance most efficiently, might involve lots of different ways of approaching the problem: from the perspective of pedagogical experts, teachers, parents, budgetary officials and last but not least, students. Each one of these groups could add a valuable perspective on the problem—not in the sense of missing information, but in the sense of—say—being able to

---

<sup>354</sup> Ober, *Democracy and Knowledge: Innovation and Learning in Classical Athens*.

<sup>355</sup> Page, “Uncertainty, Difficulty, and Complexity.”

estimate the effect of a policy intervention on a particular dimension of the problem. In contrast, chess, while being an enormously challenging game, is contained and requires but one specific skill-set rather than a diverse collection of them. The performance of a team of grandmasters (or of a super-computer) would not be improved by adding some people with fundamentally different skill-sets, like great public speakers or experts in textual exegesis.

Nevertheless, as pointed out throughout this dissertation, political problems *are* frequently complex in the relevant sense. Accordingly, it would be fair to assume that there really are benefits to a diversity of information-processing capacities within the collective decision mechanism. Coming up with good potential solutions to complex problems does not depend only on a number of people with different views all getting to express them—they must be able to evaluate and process them within the collective group. So it seems that if strictly democratic first-order decision procedures do embody this type of cognitive diversity, then our political decision-making should be organized in this way.

In a collective problem-solving effort that combines different information-processing capacities, it matters that the relevant problem-solving skills should come together at the same time (and ideally in the same place). The model of problem-solving underlying Page's theorem and Ober's model is one of a continual back-and-forth between the collective and individuals estimates. The key to good information processing is to find the relevant diversity of problem-solving skills and get them to interact in a productive way in the problem-solving process.

Does it follow that we should have democratic first-order decision procedures? First, there is the question *how* one can represent the relevant diversity within the decision process. After all, problem-solving skills are not directly observable qualities, especially not “general” problem-solving skills (i.e. problem-solving skills that are valuable across many different domains). In the school reform example above, it seems clear that someone from all the parties that are involved in the problem (parents, teachers, students, experts) should be represented, and that this would ensure the relevant diversity. However, generalizing this example may not be too easy. Depending on the *problem at hand*, the relevant diversity of heuristic tools may be highly concentrated, perhaps among a group of particularly open-minded and/or brilliant thinkers.

For instance, assume (not entirely unrealistically) that across the population nearly everyone tends to use essentially the same kind of predictive model (heuristic) in their everyday lives. You could think of the widespread mental shortcut of finding a pattern in past events that conforms to a salient narrative, and grounding our decisions on an extrapolation of that narrative.<sup>356</sup> The only ones using different models are professional forecasters in the media or in academia.<sup>357</sup> If we are looking to employ a diversity of *predictive models* in our collective decision-making process, perhaps we need only to draw from the small group of professional modelers? In that case, we do need *some* diversity, but only diversity within some clearly circumscribed group.

---

<sup>356</sup> The almost universal tendency of individuals to ascribe narrative patterns on the past forms the core of the already cited Watts, *Everything Is Obvious: Once You Know the Answer*.

<sup>357</sup> This difference maps neatly on the dispute between Nate Silver and traditional “pundits” that became salient in the last weeks of the 2012 American Presidential Election. See also Silver, *The Signal and the Noise*.



Indeed, there is no reason why it should not be a single person who has all the relevant skills in her head (or laptop).<sup>358</sup> An open-minded professional statistician, for example, may be aware of quite a wide diversity of possible predictive models, and may even employ them himself. There is no particular reason why he could represent only one (or even only one type) of approach to the problem. When Page writes that “Crowds are not wise, but crowds of models are,”<sup>359</sup> this seems to me to give rise to the question why there can’t be crowd of models in a single person’s head, let alone in his computer or in a small group of so-called experts. Acquiring a sophisticated set of cognitive tools or problem-solving strategies depends on receiving adequate training as well as life experience. Creating a diversity of useful cognitive abilities in a given population may actually require quite a high level of universal skills, a basis on which one can acquire personal cognitive tools. This again, may not be such a widely shared attribute.<sup>360</sup>

In other words, it does not seem obvious to me why in such cases the presence of cognitive diversity requires the physical presence of many people, each holding, as it were, “one unit” of cognitive diversity.<sup>361</sup> This might not necessarily be terribly realistic, but it should be clear that endorsing Page-type diversity does not imply maximizing any old measure of diversity within the decision-making collective, at least not without further

---

<sup>358</sup> Of course there may be *empirical* reasons why we think that this person does not exist, or that if she did, she would be hard to find.

<sup>359</sup> Page, *The Difference* 341.

<sup>360</sup> *ibid.*, 303.

<sup>361</sup> Page actually has an argument here: because of individual cognitive limitations, a group of people collectively can successfully estimate a model of reality that is more complex, and therefore more likely to be accurate, than the model of any single individual within the group. This discounts, of course, the possible use of technology in these kinds of decision scenarios. Furthermore, this of course still allows that the relevant diversity might well be limited to some specific group rather than the population as a whole. This latter point will be taken up below.

consideration. Whether or not the diversity of cognitive skills correlates with some other more readily observable measure of diversity is an open question.

### *Granularity*

A related problem we might call the problem of the *granularity* of the relevant diversity. Consider this example: a group of micro-biologists may be internally very diverse, in the sense of combining different methodological approaches and differences in opinion. However, from a higher-level perspective—say the perspective of a whole society—a group that includes only micro-biologists is of course anything but diverse. At which of these levels we should be selecting, obviously, depends on the problem at hand: is it a biological problem or not? Page's theorems argue that there are clear epistemic benefits to diversity at any *given* level of granularity. However, for a particular problem, not every level of granularity may be applicable.

We can see therefore that the argumentative move from diversity of processing capacities to a first-order democracy requires further argument to address this problem. There are two possible ways to make such an argument. First, one could argue that the relevant diversity for any given case (of opinion, or of cognitive skill) is *in fact* distributed very widely and evenly across the population; and that this fact makes inclusive decision-making necessary. In other words, that for political issues, we require the *coarsest* granularity possible, i.e. diversity on the society-wide level: people have to be *as different as possible*.

However, I do not see any reason why this is a cogent assumption. Many political problems are quite specific, and involve highly specialized knowledge. For example, while there is a compelling case that given its extreme complexity, economic policy should be set by a *diverse* group of economists, who employ a variety of different models and have different ways of looking at data—and especially also different political persuasions that consciously or unconsciously influence their viewpoints—it is less obvious that a more *general* diversity would be helpful. As already discussed in chapter 5, Page himself acknowledges this problem through his *Calculus Condition*, which states, roughly speaking, that added diversity only helps if the new people understand the problem well enough that they can distinguish better from worse solutions.<sup>362</sup> Is it reasonable to assume that this condition is satisfied by everyone across the board in the political sphere?

Prima facie, a more convincing way to argue would therefore be this second one: that in the political world, it is *ex ante uncertain* which skills are relevant. Political problems are often ill-defined, and it is frequently unclear even which standards of success apply. We just don't know which kind of diversity we need for any given problem—we do not know which level of granularity we are at—so that might be a prima facie reason to maximize the diversity at the highest level. There is no reason to select any particular distribution of cognitive skills over any other, and the default should be to maximize diversity as we can observe it. Hence, in order to benefit from cognitive diversity in our political decision-making, we should try and get the most diverse group of people together, i.e. we should get diversity at the coarsest level of granularity. In essence, this is the kind of argument made by Page:

---

<sup>362</sup> *ibid.*, 160.

“The results from this section do not imply that when confronted with a novel situation we should always choose a diverse collection. If we have enough information to know what drives performance, then we should select the best collection on the basis of that information. If not, and if we only get one try, then we should probably choose a diverse collection. This does not mean that diversity is always better, only that if we are not sure of what we’re doing, we should err towards greater diversity.”<sup>363</sup>

The problem with this argument is that it seems not obvious that the rational reaction to being faced with varied or unknown or uncertain problem structures is to go to the maximum level of granularity across all cases. As already mentioned above, it is not the case that the coarser the diversity, the better (a minimum condition for example is the “calculus condition”). One might of course say that adding ever more people to a decision group will at least not make it any worse, and it might “err” on the side of more diversity. However, besides the practical problem that this creation of large unwieldy collections of people who don’t understand the problem bring, we still do not know which level of granularity we should draw on in this process.

Now, if we do not know what the problem structure of future political decision situations will be, a more pragmatic view would be that what we need is an *adaptive* selection of decision-makers, not a *maximal one*. The right level of granularity of cognitive diversity, as it were, will necessarily be determined in practice.<sup>364</sup> Of course this does not work if we “only get one try,” as per the assumptions of Page’s model, but there is no reason why political activity should be understood in this way.

To illustrate this rather abstract argument, return to the house construction example. Imagine I am not sure how coarsely diverse my assembled group should be. Do

---

<sup>363</sup> Page, *Diversity and complexity* 194.

<sup>364</sup> The alert reader will immediately notice that she has read this before in this dissertation.

I want ten architects of different persuasions (a fine-grained diverse group), do I want a diversity of people involved in the construction trade (a medium-grained diverse group), or do I want an even more diverse group, say an architect, a philosopher, a financial expert, an environmental activist, etc. (a coarse diverse group)?

We can imagine situations in which each of those might be useful: if I have a big important building planned, I might actually want to get a coarse group together. At the other extreme, if I am ordering a pre-fab house and most problems are already solved, I might care only about the aesthetics and the fine-grained group of architects may come in handy. If I don't know the specificities of my problem-situation, the best approach is probably to assemble a group based on my best estimate, and then increase or decrease the granularity in response to practical results. Say I assemble a fine-grained group of architects to help with my house. Then it turns out that the house project is more difficult—the architects get stuck on a particular plumbing problem. In that case, it would pay off to increase the granularity of my group, and get a plumber involved.

Now, the point of this example is to show that *I have to select for diversity at the adequate level of granularity*. What that level is, depends on the structure of the problem at hand—and insofar that the problem I am facing is uncertain, my selection should be adaptive, not final. We cannot expect a random selection of diverse people to exhibit the cognitive diversity relevant for the task at hand. Even if, as I do not doubt, a diverse group is better at problem-solving than a uniform group, a group with the wrong kind of diversity may be even worse.

This issue is compounded by another major problem: it is difficult to make sense of the idea that we should maximize cognitive diversity *as such*. We can only maximize diversity on some measurable dimension.

Therefore, we would have to resort to some proxy dimension: ethnicity, gender, income, social class, etc. Maximizing diversity on some dimension in our decision-making group might increase the probability that the relevant diversity of cognitive skills is represented, but we can equally think of plausible cases when it might not. Remember we are working under the assumption of *uncertainty* about what will turn out to matter. For instance, assume that in a given country, certain ethnic groups are concentrated almost exclusively within a small number of large cities. In that case maximizing diversity of geographical location, which may require a drastic overrepresentation of rural areas, would probably lead to a loss of diversity in ethnic background. When, say, the correct set of economic policies depends on knowledge about the different lifestyles and preferences of rural vs. urban people, this former diversity is the one we need. If it depends on knowledge about the interests and behavioral patterns of different ethnic groups, the latter diversity should be promoted. Which one it is, depends on what we think the right economic policy might be. But this is precisely what we do not know, per assumption. The point is that whether in a given instance the *relevant* cognitive diversity is represented depends on what *relevant* means, and that depends on what the truth is that we want to find.

The best bet, it seems, is the practice of *testing* for cognitive diversity directly;<sup>365</sup> there are a number of possibilities of psychological assessments that can identify different cognitive styles. Of course this once again relies on the assumption that a given person can only have one fixed set of cognitive tools.<sup>366</sup> But even assuming that this is possible, this does suggest that the way to achieve the relevant diversity is by *active selection*.

The conclusion of this section is therefore: cognitive diversity can be expected to deliver its epistemic benefits only when it is *task-specific*. As it were, we require a specific *form* of cognitive diversity, and we need a task-specific *granularity* of diversity. And it seems that this suggests we require a mechanism that consciously *selects for* and *employs* cognitive diversity that has the right composition and the right level.

Uncertainty as to which structure of diversity is required for a given task implies that this process should be *adaptive*, i.e. able to adjust the cognitive diversity according to the practical outcomes of the process of decision-making. What this again suggests is of course that the process of selecting diverse decision-making groups should be itself *experimental*. This is of course one of the advantages of the control model of democracy: it allows for experimental institutional solutions to the first-order decision problem.

#### *How do we create diversity?*

Now let us for the moment set aside the questions of the types of diversity, and assume they can be answered. The next important question is how to get the right kind of

---

<sup>365</sup> Page, *The Difference* 360.

<sup>366</sup> Above I pointed out some problems with this assumption.

diversity represented in the decision process. For the purpose of argument, let us say that the granularity question does not arise, and we want to get the *most diverse* set of cognitive skills in our decision collective at that level.

So what we need is some way of *selecting* a set of decision-makers from the population. In keeping with the theory, selection procedures should be evaluated by the degree to which they produce a set of decision-makers with the desired cognitive diversity. For that reason, democratic mechanisms should be evaluated on exactly this basis as well. They are valuable only if we expect them to create maximal (or threshold levels of) diversity in the selection of decision-makers produced.

Now we might think that democratic elections are actually best suited for this. Democratic representation, we may think, creates a first-order decision-making collective that reflects the diversity of the general population.<sup>367</sup> However, there are a number of problems with this: even if an election produces a group of decision-makers that is approximately representative of the general population, this does of course not ensure that the cognitive diversity we want is represented.

#### *The limits of indicative representation*

First, assume for the moment that people vote for people like themselves, and assume that representative elections *do* produce an assembly that accurately reflects the cognitive diversity of the population. Of course, even under these conditions, we will not necessarily reach maximal diversity: unless the population itself is somehow already perfectly

---

<sup>367</sup> Pettit calls this idea "indicative representation," see Pettit, *On the People's Terms*.



cognitively diverse, this still does not produce maximal diversity. The largest group would be represented as the largest group in the assembly as well, and the smallest groups would possibly not be represented at all. In fact, with respect to *cognitive* diversity, as opposed to other forms of diversity, like diversity in wealth or ethnicity, this problem is exacerbated: small minorities whose only shared characteristic is a *cognitive style* are, I think it is safe to assume, unlikely to collect together and demand representation *on the basis* of that characteristic. Whether or not a certain cognitive style is represented therefore depends on whether it maps onto another salient distinctive feature.

Maximizing diversity, therefore might imply that we should *overrepresent* less popular groups, in order to get their particular cognitive style adequately represented. This, however, requires an active intervention in the democratic process: by setting quotas, by drawing constituencies in an unequitable way, or by some other measure of this kind. In other words: if we want to maximize cognitive diversity, why go the roundabout way of elections that have to be adjusted and nudged in the right direction, rather than selecting a group of decision-makers *according to its cognitive diversity*?<sup>368</sup> Especially if it is possible to test for cognitive diversity directly, this seems to be the more rational strategy.

### *The unpredictability of elections*

The second point is however, that people do not necessarily vote for people who are like them, except perhaps in a very broad sense. Consider a first-past-the-post system with

---

<sup>368</sup> See also Page, *The Difference* 367.

single-member territorial constituencies. This conventional system of representation by election in many ways will not guarantee the diversity of knowledge we need. Unless the relevant cognitive diversity should happen to be *correlated with territorial location*, there is no reason to assume that an elected representative assembly on this basis would be particularly diverse in any sense, let alone represent the diversity we need. Unless constituencies should be radically different, it seems a more plausible assumption that every constituency will send a similar type of person to the assembly; they will employ a similar criterion when choosing a representative, the type of person that seems most qualified.

Just to take an example, this was in fact one of the main concerns of the Anti-Federalists with the proposed United States Constitution. The Anti-Federalist “Brutus” puts it like this:

The great body of the yeomen of the country cannot expect any of their order in this assembly—the station will be too elevated for them to aspire to—the distance between the people and their representatives, will be so very great, that there is no probability that a farmer, however respectable, will be chosen—the mechanics of every branch, must expect to be excluded from a seat in this Body. It will and must be esteemed a station too high and exalted to be filled by any but the first men in the state, in point of fortune; so that in reality there will be no part of the people represented, but the rich, even in that branch of the legislature, which is called democratic.<sup>369</sup>

This prediction is borne out by the uniformity of background in most Western democracies employing a system of territorial representation. Almost all candidates for office are likely to have at least college education; in many countries representative tend to be relatively old, white and male; in the US Congress, almost 50 per cent of members are millionaires (while the proportion of millionaires in the US population is at around 1

---

<sup>369</sup> Alexander Hamilton, James Madison, and John Jay, *The Federalist: With Letters of Brutus*, ed. Terence Ball (Cambridge: Cambridge University Press, 2003).

per cent).<sup>370</sup> In the UK a disproportionate number of MPs (three out of ten), regardless of party or constituency, are graduates of one of only two universities;<sup>371</sup> in Germany lawyers and public sector employees are very disproportionately represented in parliament.<sup>372</sup> For the purposes of this chapter it does not matter which way the arrow of causality goes: whether voters tend to select similar kinds of people, or whether politicians self-select in such a way that the same attributes that will make you a millionaire will also tend to get you elected. The result is that similar types end up as decision-makers.

The theoretical point is easily understood: consider again the example of building a house. If I want to build a house I *consciously choose* the relevant diversity of helpers: one architect, one structural engineer, one bricklayer, one electrician etc.; imagine that instead I let all my friends each nominate one person (without allowing them to communicate). Even if my friends are a diverse bunch, why should I expect that I end up with one architect, one engineer, one bricklayer, etc., rather than with twelve architects? Ignorant of each other's choices, the rational choice for each of my friends would be to select whoever they deem most important for my house project. In contrast, the team I select exhibits the relevant diversity because it was deliberately chosen that way, for a specific purpose (namely, constructing my house).

Note that while proportional representation systems may do somewhat better than territorial representation systems, it is unlikely to make much of a difference: who gets

---

<sup>370</sup> This proportion rises to nearly two thirds when we look at only the Senate.

<sup>371</sup> <http://www.independent.co.uk/news/education/education-news/increase-in-number-of-mps-from-private-schools-1970414.html>, accessed February 2013

<sup>372</sup> [http://www.bundestag.de/bundestag/abgeordnete17/mdb\\_zahlen/Berufe.html](http://www.bundestag.de/bundestag/abgeordnete17/mdb_zahlen/Berufe.html), accessed February 2013

elected depends on who gets on party lists, and party lists are not necessarily a hallmark of diversity.<sup>373</sup> For one thing, everyone on the party list is already ideologically close (since they are members of that party), and has been socialized, like everyone else, within the various stages of the party apparatus.

The underlying problem is of course that these systems tend to select representatives according to some *common standard*, not according to a diversity of standards. A democratic system will only produce a diverse assembly if people use *different* standards of evaluation across constituencies, *and* if this difference maps onto cognitive differences.

Manual workers, homemakers, college students, let alone many people with disabilities of some kind, will remain underrepresented in such a system. One solution to this may be of course to move to a more syndicalist model of politics and design constituencies along functional rather than territorial boundaries, such that, say, manual workers elect their own candidate, and lawyers pick their own. However, it is important to note that this requires an active intervention into the democratic process, and while such a system may not be any less democratic than the one we have now, it is clear that whether or not a system produces diversity depends on the way it is deliberately set up, not only on whether it is democratic.

The simple upshot of this discussion is thus as follows: democratic systems will produce a maximally diverse body of representatives only if (a) the underlying population is already maximally diverse, and is divided into a set of constituencies that is itself a

---

<sup>373</sup> Many parties, at least those on the more progressive end of the spectrum, self-impose some quotas that are supposed to ensure diversity. Party lists may have to contain a minimum percentage of female candidates, candidates from immigrant backgrounds, etc.

maximally diverse set. That is, individual constituencies should not contain diverse populations; quite the opposite: they ought to be as different from each other as possible. The second condition is (b) the voters select people who share their cognitive approaches and not people who they believe will be the best according to some widely shared standard. And these are two pretty substantial “ifs.”

Thus, we can see, that institutional mechanisms other than the conventional representative ones, may better increase the availability of cognitive diversity within the decision process. Conventional representative elections will only *incidentally* create a diverse decision collective, and it may require intervention precisely with the goal of creating greater diversity. If we want more underrepresented people influencing the decision process, we need to disproportionately increase their chances of doing so.

Of course this is *compatible* with democracy as it is commonly understood, but given these issues it seems difficult to form a *justification* of democracy on the basis of its optimal (or adequate) creation and application of cognitive diversity. The last two sections have dealt with the problem of the nature of diversity we need, and the mechanisms by which such diversity may be created. The following section turns to the policy-process conceived as a whole, and considers at which point in this process we require the adequate diversity.

*Where does the diversity apply?*

Now, let’s complicate the picture even further. Setting aside the question of the *kind* of diversity we want, we have to face the question *where* in the first-order decision-making

process diversity is supposed to apply. Let us think of a crudely simplified model of the political decision-making process, one with only three stages: coming up with proposals, deliberating on different proposals and narrowing down the choice, and finally deciding which proposal to adopt. Assuming for the moment only these three phases, where do we want the relevant diversity?

During the proposal-creation phase the wide array of all possible solutions is brought up. In the next phase, this variety is narrowed down into a concrete legislative proposal (or perhaps a finite number of rival proposals). In the following phase a decision is made between different proposals, or if there is only one, a decision whether it should be adopted or not. This is usually done by some sort of vote preceded by discussion of the respective merits of the proposals (in so far that this discussion does not take place in the previous phase). After the decision the new rule or policy is enacted.

So where does the epistemic benefit of diversity come in? It seems to me that most clearly, diversity of available information is required at the *creating proposals* phase of the decision process, and diversity of information-processing capacity is required at the *deliberative narrowing down* phase. These are the phases of the political process that correspond most obviously to the idea of collective problem-solving activity. What is key is that our diversity mechanism applies *before the act of decision itself*, and consists both in the *creation of diverse possible solutions and the initial evaluation and (deliberative) contestation of these solutions*. Here the diversity of problem-solving skills are applied and here is where their epistemic benefits are realized. Roughly speaking, we want input from a diversity of people when we have to come up with different possible ideas, and then when we want to

discuss whether these ideas are any good. This corresponds to what above is called the *variation* function within the experimental model of policy-making.

This seems somewhat plausible. Both in the creation of different possibilities, and in the deliberative contestation there are clear benefits to getting new cognitive perspectives involved: coming up with an idea nobody has thought of, or with an implication of a policy that the collective was previously unaware of—these things require cognitive diversity.

What about the *decision* phase? Think of the up-or-down vote in parliament after a final proposal (or set of proposals) has been decided upon. Here, the preferences of the individual members are aggregated, and the choice of the majority gets adopted. Does diversity have any epistemic benefits here?

This seems much less obvious. Given that—epistemically speaking—the task of the collective at this point is *not* to gradually work towards a solution, but to make a one-off judgment whether the final proposal will actually solve the problem at hand or not, this would mean that we have to trust that diverse assemblies are more likely to adopt rather than reject a good proposal once it has been developed. That is their role, from an epistemic standpoint. Not only would they have to identify a good proposal once they see it, they also have to act sincerely and vote according to their honest assessment. This is quite a different requirement than cognitive diversity: what it requires is *problem-focused*, or if you like, *common good-oriented* behaviour. At the decision phase the key is that representatives really do have the interests of the population in mind when deciding whether to adopt a policy.

This means that at this stage cognitive diversity does not necessarily confer an advantage. The cognitive and informational diversity (should) have had their input already at the design phase. Recall that what cognitive diversity is supposed to do is to collaboratively find creative solutions; and this works by everyone contributing their individual models of the world, such that the collective model of the world is more likely to be accurate. In contrast, in the decision phase, where everyone votes independently, the proposal will be evaluated with everyone's individual model in mind.

Indeed, having cognitive diversity at the decision-phase may even be counter-productive, namely if cognitive diversity is correlated with diversity of fundamental preferences; this would make strategic misrepresentation of judgments among the representatives more likely. So far throughout this section we have assumed that diversity has no drawbacks in practice—however, we can easily imagine that a multiplication of cross-cutting cognitive perspectives within a political decision collective may complicate the finding of rational solutions.

Thus, it seems that cognitive diversity is required at the early, problem-solving focused phases of the policy process, while the judgment phase requires mainly a unity of purpose. Both Page and Ober model the political process as a diachronic, problem-focused activity rather than a purely aggregative procedure.

Now, we may think that if the same collective body creates proposals, narrows them down, and takes the final decision, this still is a strong argument in favor of creating a diverse collective first-order decision-maker. However, as the discussion of the control model has already suggested above, there is no particular reason why it must be the *same*



*collective agent* that is active in those different phases of policy-making. Creating proposals, narrowing down options, and final decision could be undertaken by different groups. Before this is dismissed as a merely theoretical possibility, note that in actual democracies, to varying degrees *this is the usual practice already*. In the United States system, proposals come from the floor of Congress, are then deliberated upon in the committee system, and are then referred back to the plenary of both houses. In parliamentary systems, such as the UK and Germany the creating proposals and narrowing down phases often take place *within* government or the coalition itself—either within the cabinet or within the parliamentary parties, and only the decision phase is taken by the assembly (and even this is sometimes a formality if government has a solid majority). Furthermore, who is and is not consulted in those decision processes is highly fluid: in addition to elected representatives, there may be stakeholders, experts, bureaucrats and lobbyists involved. Decisions are made in networks with a variable membership.<sup>374</sup>

What is striking is that, according to the diversity analysis, these parliamentary systems of course have it exactly backwards: the least diverse and most hierarchically organized bodies operate at the proposal creation and deliberation phases, and the purportedly most diverse body (the plenary of the legislature) operates only at the decision phase. Of course, governments *may* include diverse mechanisms in the relevant phases of the policy process, but there is no institutional requirement or incentive for them to do so. Governments may consult dissenting voices, but they may not do so: they may just rely on the advice of their ideological peers, or on pressure groups that happen to gain access to government.

---

<sup>374</sup> Chalmers, *Reforming Democracies*.

Therefore, it seems that to some extent there is already a division of labor between different subsets of the legislative branch or between different political agents with respect to the policy process. This again suggests that we are in particular need of improving cognitive diversity within the fora in which the true deliberation on proposals is going on. Increasing cognitive diversity among deputies does not, under realistic conditions, necessarily guarantee that the diversity will have an effect on the crucial phases of the policy process.

Indeed many political systems amend their policy process with precisely such mechanisms of consultation that increase the cognitive diversity of the decision-making mechanism. Committee hearings, either of experts or of minority representatives are common to many representative systems. In Germany, the parliament may also institute so-called *Enquetekommissionen*, committees staffed partly by elected representatives and partly by a variety of appointed experts. In the European Parliament, committees may contract up to two research institutes and/or consultants for each legislative project, who then supply their cognitive skills to the committee (while not receiving any voting rights, of course).

These are just some illustrations of how cognitive diversity can be added to the decision procedure *beyond* the level represented in parliament. A key problem of this is of course the possibility of abusing these diversity-enhancing consultation mechanisms: the majority faction may staff such a commission only with so-called experts that do nothing but confirm the majority viewpoint. The appointment to consultatory commissions may, as is not unlikely, devolve into another arena of partisan conflict. For that reason, we have to pay attention to the design of the embedding institutions as well.

Given the discussion of diversity mechanisms above, we can conclude that (1) the *creation of proposals* requires a diversity of information, which in turn requires both an environment where a diversity of potential solutions can be created *and* an institutional mechanism that can access of this networked information; (2) the *narrowing down* phase requires information *processing* and therefore may benefit from diversity of cognitive skills; (3) the decision phase, which aggregates individual evaluations of proposals above all requires sincerity and common good-oriented behaviour.

Now given the discussions in the foregoing sections, it is an open question whether democratic mechanisms will be ideal at creating diversity at phases (1) and (2). Depending on the particular situation, a purposefully selected group of diverse representatives may be much better at bringing problem-solving to bear.

Combining the conclusions from the three sections, therefore, we can see that cognitive diversity has only an incidental relationship to democratic mechanisms. Of course, as already mentioned above, democracy may be perfectly *compatible* with a system that optimizes the use of cognitive diversity in the proposal and deliberation phases of political decision-making, but in itself it may not actually bring about such an effect. The obvious alternative is a system that actively selects *for* cognitive diversity in a *task-specific* way, and employs this diversity in the most problem-focused phases of the policy process. Imagine, for instance, a parliamentary system which for every political problem creates a special committee to deal with the issue. This committee is not staffed, as is the usual practice, proportional to parliamentary representation, but according to its cognitive diversity in a task-specific way. Furthermore, the membership of the committee is flexible and adaptive, and can be adjusted depending on how the problem-solving process is

working. They are tasked with designing a proposal, which then is submitted for approval to the full democratic assembly.

However, I expect there will immediately be a question regarding this hypothetical system: who is to ensure that these committees are selected properly? Especially if their membership is flexible and adaptive, who is to guarantee that the people in charge are not packing the committees to their own advantage? We need, as it were, a *second-order mechanism*, whose task is not to solve political problems, but to keep those people tasked with problem solving on track.

Of course, this is precisely the role played by democratic institutions in the control model. We can see therefore that the cognitive diversity argument is compatible with experimental democracy. Beyond that, however, the control model gives us a clearer picture of how this diversity can actually really be brought to bear on the problem-solving process itself.

#### 4. Making Feedback Effective

Now we have seen some suggestions of how the *identification* and the *variation* functions of experimental democracy could be realized effectively. Key aspects of an experimental approach to problem solving is to bring the problem itself onto the agenda and to come up with some problem approaches that are then implemented in a controlled way. Effective experimental methodology, however, also requires an ex post *evaluation* of the

success or failure of the experiment, and then *adaptation* to these results: proceeding further in cases of success, revision or reversal in the case of failure.

In other words, the experimental method requires adequate *feedback* about whether the experiment (in policy or institutional choice) works or not. Recall the pragmatist definition of a “problem,” as defined in chapter 4 above: The important question is whether or not the clash of ends-in-view within the concrete experience of people has been resolved or not. This information therefore must get to decision-makers, and, equally importantly, they must have an incentive to adapt accordingly: to attempt to reverse the policy if it did not resolve anything (or if it has made things worse), or to further pursue the approach in cases of success. In other words, to change their mind in response to the new information. Of course, decision-makers who do not want to adapt will be removed in this model: the system as a whole can adapt even if none of the units will. But both of these ways rely on an effective mechanism that causally links the effect of the policy back to the incentive structure faced by policy-makers in a feedback loop.

In the control model, this mechanism is realized in elections of representatives. Therefore, this issue calls for more investigation: can elections really fulfill this function? Which framework rules are necessary such that elections *are* effective feedback? This is the topic of this section.

I assume for the purposes of this discussion that people are competent to assess whenever the problems in their experience are actually resolved and when they remain unresolved.<sup>375</sup> Unless they have countervailing incentives, I also assume that they will

---

<sup>375</sup> This assumption was discussed in chapter 5.

want to express their satisfaction or dissatisfaction with their general life situation through the vote (but not that they will necessarily take a common-good perspective).

On the basis of these assumptions, I focus on two key obstacles that could prevent elections from being effective feedback mechanisms: (1) a failure of accountability in the sense of a failure on the part of the voters to connect the effects of policy with the policy itself, (2) and the paradoxes of social choice. In the case of (1) voters would misidentify whether a policy (and thereby an individual decision-maker or party) have been successful or not. The main reason why this may happen is *muddled accountability*, and as I will argue the appropriate response is a *comprehensibility* requirement. Point (2) refers to notions from social choice theory like the Condorcet Paradox, Arrow's Theorem, and the Gibbard-Satterthwaite theorem. Together, they suggest that all collective decisions may to some extent be arbitrary artifacts of the decision rule. Thus, even if people *do* correctly identify the success or failure of a policy, the process by which this collective judgment is transmitted may distort the collective decision such that accountability is misdirected and false incentives are set. This would of course also derail the adaptation function. As I argue, this second problem calls for a more fundamental institutional change: if we want the electorate to give informative feedback, we should ask them for more information than just their vote. I will address these two problems in turn in this last section.

### *Comprehensibility*

The right solution to the problem of muddled accountability is often taken to be *transparency*.<sup>376</sup> When voters can observe who among the deputies voted for which policy—and if the information is in the public record—they can establish the link necessary to hold them accountable.

The discussion of these issues therefore tends to focus on weighing the benefits and drawbacks of publicity in decision-making. Publicity facilitates accountability in the sense just mentioned. It may also encourage decision-makers to present publically acceptable reasons for the proposals they put forward in debate. This may focus decision-makers on addressing the problems at hand, and at least publicly declaring good faith attempts to resolve the issues brought onto the agenda.<sup>377</sup> In conditions of secrecy where there is no such “civilizing” framework, naked self-interest can enter the policy-making process more easily, and the possibility of covertly dismissing problems brought forward for no substantive reason is possible (see section 2 above).

On the other hand, it may encourage grandstanding and playing “to the audience” on the part of political decision-makers. Forcing decision-makers to conduct debates and to vote on proposals in public may prevent them from reaching bargains, compromises and deals that they would be able to reach in more secluded settings. Sometimes the subject-matter of the policy decision may also contain sensitive information, such that publicity would defeat the purpose of the decision in the first place.

---

<sup>376</sup> See for instance Vermeule, *Mechanisms of Democracy: Institutional Design Writ Small* ch. 6; Elster, *Securities against Misrule*; Rahul Sagar, *Secrets and Leaks: The Dilemma of State Secrecy* (Princeton: Princeton University Press, 2013).

<sup>377</sup> see also Green, *The Eyes of the People : Democracy in an Age of Spectatorship*.

Sensitive information about national security is the key example here but there issues beyond that as well. This is complicated by the fact that an efficient oversight over which decisions should be disclosed and which should not is difficult, if not conceptually incoherent.<sup>378</sup>

One solution that has been discussed here is *delayed disclosure*, or *ex ante* secrecy paired with *ex post* publicity.<sup>379</sup> This may minimize the immediate risks as consequences of the publicity of the decision making process. At the same time, the shadow of future disclosure may discipline the candidates in the same way as direct publicity. Furthermore, *feedback* of course requires publicity only at the point of evaluation (that is, at the point of an election). This is because citizens do not constantly have to keep track of a decision-making procedure during the proposal-generation phases. They do not necessary have to know how a specific policy was made at the time, as long as they can rely on the ex post publicity that is required to connect their evaluation of the success and failure of a policy to concrete voting actions.

Thus, I shall assume that there may be an institutional solution to this transparency issue. The problem of connecting concrete *policies* to concrete *political actors* (representatives, factions, or parties) could possibly be resolved through appropriate framework rules.

However, there is another key problem before this: voters may have difficulty connecting the *outcome* that they observe with the *policy* that brought it about. Faced with

---

<sup>378</sup> Sagar, *Secrets and Leaks: The Dilemma of State Secrecy*.

<sup>379</sup> Vermeule, *Mechanisms of Democracy: Institutional Design Writ Small* ch. 6; Elster, *Securities against Misrule* ch. 2-3.



extremely complex and uncertain problem contexts—as is of course the basic general background assumption of this whole dissertation—it may in principle be difficult to identify which policy (if any) was causally responsible for the problem-solution (or failure to solve the problem). This is compounded by the fact that without expert knowledge and/or legal training it is frequently not easily comprehensible what a given policy actually does (many policy provisions run into hundreds or thousands of pages). More problematic than the particular mechanisms of a given law is however, that it is often unclear what the policy is actually *intended* to achieve; in the language of pragmatic problem-solving, it is frequently unclear *which problem* the policy actually addresses, and how we would know when it does.

This knowledge, however, is essential to make an experimental problem-solving methodology work properly. In the language of experimentation, we need to have a clear, comprehensible *hypothesis*, on the basis of which we can derive expectations of what we should observe in the case the hypothesis is true (or false). Only if we can decide on the basis of observation or more broadly, “experience,” whether the experiment has caused the expected outcome or not, can we infer anything about the hypothesis from it.

Applying this principle to the model of democracy, therefore, this means that a functioning experimental democracy needs what we may call strict *comprehensibility* requirements on the actions of the political problem-solving process. In particular, if we understand policies as experimental interventions, they must be accompanied by a public declaration of the “hypothesis,” i.e. the problem they are intended to solve, and how they attempt to solve it. Furthermore, the “expected results” must be specified. That is, any policy must specify standards by which its success could be measured; non-abstract

standards by which voters may determine whether the goal was reached or not. Ideally, policies will clearly define their own success standards. A rule that this is required from any policy proposal is very easily implemented, but would have a great effect in terms of the adaptability of the system as a whole.

Now this does not solve the causal complexity problem in a fully satisfying way: we cannot infer in any meaningful sense whether the policy really caused the outcome unless we have a fully specified experimental situation: randomization, control groups, constant environmental conditions, etc. But as already discussed in chapter 4 above, finding the “truth” about these connections is not the primary task of the political process. That is adaptive improvement. The task, as discussed in chapter 3, is *problem-solving*. As such, we are mainly interested in whether a problem has been solved or not; and if the election process results in the occasional false positive (a politician gets lucky, and a problem goes away on its own), is less problematic.

The second great advantage of comprehensibility requirements is that not only do they enable the ex post evaluation of whether the policy has worked or not, they also enable us to *subject the success criterion itself* to scrutiny. Assume for example that a given policy has as a stated goal the improvement of economic well-being of all classes of society, and that it specifies as its success standard a rise in GDP of at least, say 1.5 per cent. Then, at the next election, we can of course evaluate whether this standard has been reached. Moreover, however, we can also evaluate whether that standard based on a rise in GDP was a good criterion for the original problem that motivated the policy. Therefore comprehensibility is necessary for this reflexivity function of democratic decision-making as well.

### *Social Choice*

So there may be institutional options to possibly mitigate the problem of connecting policies to outcomes. A more weighty problem however arises from social choice theory. One of the main conclusions of the different theorems here are that we cannot infer any particular meaning from any collective choice at all.<sup>380</sup> The feedback function, however, depends on the notion that we can *interpret the collective judgment expressed in the vote* as a bona fide judgment about the merits of certain policies—a good faith evaluation of legislative success or failure.

There are a number of “impossibility” results with respect to collective judgments. The problem of possible Condorcet Cycles means that collective preferences may be intransitive, and that the result of a collective majority vote therefore may depend completely on the order in which the alternatives are voted upon. If this is a possibility, then of course the judgment may be determined only by the agenda setter, and we cannot interpret in a meaningful sense as a judgment on the problem-solving efforts of the policy-makers.

Arrow’s Impossibility Theorem suggests that there exists no collective choice procedure that satisfies four normative desiderata of such a procedure: unrestricted domain (U) (i.e. that all possible profile inputs are allowable), independence of irrelevant alternatives (IIA) (that the pairwise ranking of two options does not depend on the presence or absence of a third alternative), non-dictatorship (N) (the procedure may not depend only on the choice of a single person), and Pareto efficiency (P). These are not

---

<sup>380</sup> See for instance the classic discussion in Riker, *Liberalism against Populism : A Confrontation Between the Theory of Democracy and the Theory of Social Choice*.

only generally desirable attributes of a collective decision mechanism, with respect to the experimental model they are essential. Let me just illustrate this with the IIA condition.

If you recall the discussion in chapter 3, section 2 and 3, one of the key ideas of how to understand problem-solving under disagreement was that a procedure should identify the possibility of *partial* progress even if ultimate value disagreement remains unresolved. IIA complicates this substantially. If the collective evaluation of two options (say, the status quo and a marginal improvement) does not only depend on their relative evaluation, but on the presence or absence of other options (say, conflicting ultimate value commitments) this may mean that the collective judgment really is not on the merits of what is to be evaluated. The consequence could be that we fail to reach the maximal set.

Finally, the Gibbard-Satterthwaite theorem suggests that there is no non-dictatorial, inclusive collective choice rule that is *strategy-proof*. This means that there exists no such aggregation rule such that people may not better off by strategically misrepresenting their actual views. Again, this spells trouble for the feedback function of the experimental democracy model. If under every voting rule voters have an incentive to strategically misrepresent their actual evaluations, then how can we trust that the outcome of the vote is in any way an indication of the success or failure of problem-solving attempts? The findings from social choice theory are therefore highly problematic for the experimental model. There may be some hope in arguments that the paradoxa occur less often than one may think,<sup>381</sup> but this puts into question how *robust* in the public reason sense any justification of democracy on these terms can be.

---

<sup>381</sup> See especially Gerry Mackie, *Democracy Defended* (Cambridge: Cambridge University Press, 2003).

### *Voting and Information*

Somewhat unrelated to the social choice paradoxa, there is another key problem: the typical vote carries very little information. Especially in systems where every voter only expresses a single preference for a candidate (rather than ranking the entire set of candidates), we cannot infer much from this. Does this express approval of that candidate and disapproval of all the others? Perhaps the one is just the “best of a bad bunch?” Or perhaps the voter thinks highly of a number of candidates or parties, but since she is forced to pick one, makes the choice at random from the “maximal set?” Two people may vote for the same candidate, but evaluate her very differently. We do not know, the preference vote is not sensitive to this information.

The voting process in the traditional model therefore represents an informational bottleneck. Think, for a moment, of an hourglass: at the top voters have complex and subtle opinions about the candidates or parties, about the success and failure of their policies, and the directions in which they wish political activity to go. At the bottom, we require complex and detailed feedback in order to make the system more adaptive, to identify problems, and to correct mistakes and change wrong directions, if necessary. But the information gets narrowed down into a single preference vote (through the narrow point in the center of the hourglass, so to speak), from which we have to infer what voters think in all its details.

Therefore traditional voting not only cannot be relied on to represent bona fide evaluations of the problem solving performance of the people in charge. Even if that were

not a problem, we cannot learn much from votes traditionally understood. Fortunately, there may be a solution to both.

### *Majority Judgment*

Michel Balinski and Rida Laraki have only recently developed an important alternative to the traditional model of voting.<sup>382</sup> One of the main motivations is the simple observation that if we want more detailed information from our voters, why not ask them to provide it? There is no particular reason why the whole complex evaluation should have been expressed in a single preference vote (for candidate or party). For that reason, Balinski and Laraki argue that democratic elections ought to be modeled on other processes of collective *evaluation* or *judgment*: wine tasting for instance, or ice skating competitions.

In these fields, as well as many others, judges do not express their one preference. Instead, they give precise *grades* to all the candidates—all the wines get a specific numerical grade from each of the jurors; and the same is true of ice skating competitors. Now Balinski and Laraki extend this logic to the political sphere as well. They propose that instead of *voting* for their one preferred candidate or expressing their ranking of the available candidates, we should replace this system with one where voters have the opportunity to give ordinal grades to all of the candidates standing for election. The authors call this system *majority judgment*.

Without going too deep into this complex, but surprisingly intuitive proposal, let me just outline a few key features. First, voters do not just choose their most preferred

---

<sup>382</sup> Balinski and Laraki, *Majority Judgment : Measuring, Ranking, and Electing*.

candidate, but give an evaluation on a seven-point scale for each of the candidates (failure to give a grade is counted as the lowest grade). This has a few very useful implications; voters can express in more subtle ways the differences: for instance they can distinguish between candidates that they generally approve of, by giving them slightly different, but positive grades. At the same time, extremist candidates can be rejected more forcefully by getting assigned the lowest grade. Thus, this method of aggregating judgments is sensitive (to some extent) to the intensity of positive or negative preferences voters have over candidates.

Second, the winner of the majority judgment is the person with the highest *majority grade*. This is the grade that fulfills two conditions: a majority of voters has given *at least* that grade or higher, *and* a majority of voters have given that grade *or lower*. Thus, the grade is not the one given by the majority or plurality of voters, and neither is it the average of all grades (such as in ice skating competitions). The majority grade has the property that if it were lowered, a majority would prefer it to be higher, and if it were raised, a majority of voters would prefer it to be lower. The outcome of the election is then simply the candidate with the highest majority grade (subject to a number of tie-breaking rules, which I will not get into here).

Now this once again has a number of extremely useful consequences. First, this method is *strategy proof in grading*. That means that voters have *no* strategic incentive to misrepresent their grade in order to influence the grade of some particular candidate. Assume I think Candidate Anne deserves the second highest grade (call this a “B”)<sup>383</sup> I

---

<sup>383</sup> Balinski and Laraki have descriptive categories here, from “excellent” to “to reject,” but I am simplifying for the sake of example.

want to make sure she gets the B, so can I raise her grade by giving her an A, misrepresenting my preferences? The answer is no: my best strategy is still to vote honestly and give her a B. If her majority grade is already B, my A would not raise it. If her majority grade is lower than B, a vote of B and A will have the same effect. Majority judgment therefore is immune to a strategic manipulability in this sense.

This method of judgment aggregation is also *partially strategy proof in ranking*. That is, when I want my candidate not to have the highest grade, but to *win* ahead of the others, I may have an incentive to misrepresent my grade and adjust it upwards to improve the relative chances of my preferred candidate. However, the method is partially strategy proof, because if I insincerely raise the grade for one candidate, I have no incentive to then also lower the grade for their opponents. Assume once again that I honestly think Candidate Anne deserves a B. I want to make sure, though, that she comes out ahead of Bob, whom I consider a C. In that case it may pay off for me to insincerely give Anne an A.<sup>384</sup> But in that case, it would *not* also pay off for me to lower my evaluation of Bob to a D or F. The same holds the other way around, as well.

This method therefore discourages strategic misrepresentation of their true evaluations on the part of voters. Voters may, of course, also give the *same* grade to two or more candidates—they may, for instance, reject all of the candidates, maybe expressing their wish to “throw all of the rascals out.”

Another important property of this method of aggregation is that the intensity of preferences is to some extent included in the overall grade, but the system does not

---

<sup>384</sup> The reason lies in the tie-breaker rules for when candidates have the same majority-grade.



succumb to the problem those that average cardinal scores (such as the Borda count, or the ice skating example). For example: a candidate who has vocal minority support (they give him an A), but is despised by a majority (they give him the lowest possible grade—think of extreme right-wing candidates for example), would be assigned the lowest possible majority grade. If grades were averaged, the candidate could of course get a much higher average grade due to the high grades from the minority supporters. This system which is sensitive to these subtle attitude differences therefore is a huge improvement from the perspective of the vote's feedback function. Beyond that, majority judgment is *not* subject to Arrow's impossibility theorem, and therefore importantly preserves IIA and Pareto efficiency as well.<sup>385</sup>

Finally, the majority judgment method encourages voters to give *absolute* rather than *relative* evaluations of candidates or parties, which is exactly what we want in the experimental model. This is illustrated by IIA: the collective evaluation of, say, Al Gore's suitability as a candidate should not be influenced by whether Ralph Nader decides to run for president or not.

The key upshot of this discussion is that if we want a reliable public evaluation of candidates or parties, we should just *ask citizens for that evaluation*, rather than have them convert their evaluation into a strict ranking or a single preference vote. And the foregoing discussion indicates that this is very well possible.

One problem with implementing a system of majority judgment is of course that we have to be sure that there is a common language of evaluation, i.e. that an "A" or a

---

<sup>385</sup> Balinski and Laraki, *Majority Judgment*.

grade of “poor” means the same to everyone. The majority-grade system however means that this is less of a problem than with methods that average grades. However, at the same time, the language must include sufficient detail and gradations that it improves on the single preference vote. Ideally, as it were, we would interview every voter individually, but this is infeasible. Between that and the extremely blunt instrument of the single preference vote, we have to find a midway point.

While Balinski and Laraki’s view seems to me a definite improvement over the present practice of voting, especially from the standpoint of experimental democracy I do not want to go any further in defending it here. What the discussion in this chapter 6 was supposed to show is *how* we can evaluate different aspects of democratic decision-making. Experimental democracy gives us a perspective both on *what we need*, and on *how we should see* different reform proposals. Whenever we are faced with a question about the normative value of, say, different campaign finance regimes, or the question whether to have more or fewer direct-democratic ballot initiatives, we can ask whether they would contribute to or contradict the identification of political problems, the creation of variation in possible solutions, and the feedback function that enables diachronic adaptation.

## 5. Putting it All Together

This discussion of concrete institutions concludes the long and winding argument that led us here. The dissertation began with the basic question if there is a way to distinguish different forms of democracy in terms of the quality of the outcomes they produce. However, the problem was that we had no incontrovertible standard for what this quality is.

I hope the foregoing six chapters have illuminated this question somewhat. Let me just take this last section to quickly walk through the argument as a whole, once more. Basically, the whole project establishes a number of propositions, that, taken together, comprise the case for experimental democracy as a legitimate and normatively valuable form of democratic rule. Let me address these propositions now, very briefly.

- (1) The quality of political outcomes is a necessary element in our evaluations of political regimes, and thereby also in any plausible theory of political legitimacy.

This point was established in chapter 2. I have argued against pure proceduralism as a plausible theory of political legitimacy, and against theories that establish democracy as a default decision-making procedure under uncertainty.

- (2) Political problems are in many cases subject to fundamental *uncertainty*, and this implies that any valuable political procedure must be expected to deal with political problems despite this uncertainty.

- (3) This means that any justification of legitimate political authority should be *robust*, that is, it should hold up under a wide range of reasonable assumptions about what a political solution may be, or about the capacities and motivations of people.

These two points were the subject of chapters 2 and 3. They establish the normative principle that I call *pragmatic robust instrumentalism*:

- (4) It is in everyone's reasonable interest that there should be a political system that is forced to and is robustly capable of solving common political problems, as they arise, under conditions of uncertainty, and whatever the solution may be (compared to its alternatives).

On the basis of this standard I establish, throughout chapter 3 and 4, that we can robustly justify a political decision-making method that is essentially *experimental*.

- (5) The most reasonable method of problem-solving under uncertainty is an *experimental* one. For that reason, an experimental mode of political decision-making can be justified.

The arguments supporting this point are found in the latter sections of chapter 3 and in chapter 4. They are expressed in connection with a *pragmatic* understanding of the epistemic task of political decision-making. Chapter 4 also addresses a number of *objections* to this account of political legitimacy: that one cannot experiment with values, or that coercive experimentation is ethically problematic. In response, I outline the pragmatic theory of problem-solving as the basis of normatively acceptable inquiry, and compare the experimental method of policy-making to its alternatives. Chapter 5, then, establishes

the *control* model of democracy as the embodiment of the experimental mode of policy-making.

- (6) Since we cannot robustly say which first-order decision-making mechanism (democratic or not) will be good at problem-solving for specific instances, this choice should itself be experimental.
- (7) This experimental mode of first-order decision-making must be subject to effective democratic *control*, and therefore must be embedded within second-order democratic institutions if it is to function properly.

Chapter 5 makes both of these points, first in terms of a critique of first-order models of democracy, and then in a positive sense, supporting the control model. Chapter 6 finally moves on to the specific functional requirements of experimental democracy, and discusses with the use of some examples how they may be addressed through institutional design.

- (8) The three key functions of an experimental model of policy-making—identification, variation, and feedback—can be fulfilled through the control model of democratic politics. This may, however, involve some specific institutional forms.

This dissertation therefore gives us a handle on how we should understand institutional choice *within* democratic systems in situations where we are faced with extensive complexity and uncertainty. At the same time, it indicates *where* the normative value of democracy over other systems really lies. Frequently, we cannot give conclusive answers to political problems, but we can see that democracy is capable of dealing

precisely with this situation. After all, resolving all problems ourselves cannot be the task of political theory. In the words of Dewey, “It is not the business of political philosophy and science to determine what the state in general should or must be. What they may do is to aid in creation of methods that experimentation may go on less blindly, less at the mercy of accident, more intelligently, so that men may learn from their errors and profit by their success.”<sup>386</sup>

---

<sup>386</sup> Dewey, *The Public and its Problems* 34.

## BIBLIOGRAPHY

- Allard-Tremblay, Yann. "The Epistemic Edge of Majority Voting Over Lottery Voting." *Res Publica* 18, no. 3 (2012): 207–223.
- Anderson, Elizabeth. "An Epistemic Defense of Democracy: David Estlund's Democratic Authority." *Episteme* 5, no. 1 (2008): 129–139.
- Anderson, Elizabeth. "Democracy: Instrumental vs. Non-Instrumental Value." In *Contemporary Debates in Political Philosophy*, edited by Thomas Christiano and John Christman, 213–227, Oxford: Wiley-Blackwell, 2009.
- Anderson, Elizabeth. "The Epistemology of Democracy." *Episteme* 3, no. 1 (2006): 8–22.
- Ansell, Christopher K, *Pragmatist Democracy: Evolutionary Learning as Public Philosophy* (New York: Oxford University Press, 2011)
- Ansell, Christopher K, "What Is a 'Democratic Experiment'?", *Contemporary Pragmatism* 9, no. 2 (2012)
- Arneson, Richard J. "Democracy Is Not Intrinsically Just." In *Justice and Democracy: Essays for Brian Barry*, edited by Keith Dowding, Robert E Goodin, and Carole Pateman, 40–58, Cambridge: Cambridge University Press, 2004.
- Arneson, Richard J. "The Supposed Right to a Democratic Say." In *Contemporary Debates in Political Philosophy*, edited by Thomas Christiano and John Christman, 197–212, Oxford: Wiley-Blackwell, 2009.
- Attari, Shahzeen Z, Michael L DeKay, Cliff I Davidson, and Wändi Bruine de Bruin. "Public Perceptions of Energy Consumption and Savings." *Proceedings of the National Academy of Sciences of the United States* (2010).
- Axelrod, Robert M, and Michael D Cohen. *Harnessing Complexity : Organizational Implications of a Scientific Frontier*, New York: Free Press, 1999.
- Bak, Per, and Kan Chan. "Self-Organized Criticality." *Scientific American* 264, no. 1 (January 1991): 46–53.
- Balinski, Michel, and Rida Laraki. *Majority Judgment : Measuring, Ranking, and Electing*, Cambridge, MA: MIT Press, 2010.
- Banerjee, Abhijit V, and Esther Duflo. *Poor Economics : a Radical Rethinking of the Way to Fight Global Poverty*. 1st ed., New York: PublicAffairs, 2011.
- Banerjee, Abhijit V, and Esther Duflo. "The Experimental Approach to Development

Economics,” 2008.

- Barry, Brian. *Political Argument*, London: Routledge & Kegan Paul, 1965.
- Beitz, Charles R. *Political Equality*, Princeton, NJ: Princeton University Press, 1989.
- Berlin, Isaiah. “John Stuart Mill and the Ends of Life.” In *Liberty*, edited by Henry Hardy and Ian Harris, Oxford: Oxford University Press, 2002.
- Bostrom, Nick, and Toby Ord. “The Reversal Test: Eliminating Status Quo Bias in Applied Ethics.” *Ethics* 116, no. 4 (2006): 656–679.
- Brennan, Jason. “Polluting the Polls: When Citizens Should Not Vote.” *Australasian Journal of Philosophy* 87, no. 4 (2009): 535–549.
- Brennan, Jason. *The Ethics of Voting*, Princeton, NJ: Princeton University Press, 2011.
- Brennan, Jason. “The Right to a Competent Electorate.” *The Philosophical Quarterly* (April 2011).
- Brighouse, Harry, and Marc Fleurbaey. “Democracy and Proportionality.” *The Journal of Political Philosophy* 18, no. 2 (2010): 137–155.
- Buchanan, Mark. *Ubiquity : Why Catastrophes Happen*, London: Weidenfeld & Nicolson, 2000.
- Caplan, Bryan. *The Myth of the Rational Voter : Why Democracies Choose Bad Policies*, Princeton NJ: Princeton University Press, 2007.
- Cartwright, Nancy. “What Are Randomised Controlled Trials Good for?.” *Philosophical Studies* 147, no. 1 (2010): 59–70.
- Cartwright, Nancy, and Jeremy Hardie. *Evidence-Based Policy: a Practical Guide to Doing It Better*, Oxford: Oxford University Press, 2012.
- Chalmers, Douglas A. *Reforming Democracies*, New York: Columbia University Press, 2013.
- Christian, Leah. *Continuing Partisan Divide in Views of Global Warming*, Washington, DC: Pew Research Center for the People and the Press, April 2, 2013.
- Christiano, Thomas. “Debate: Estlund on Democratic Authority.” *The Journal of Political Philosophy* 17, no. 2 (2009): 228–240.
- Christiano, Thomas. *The Constitution of Equality : Democratic Authority and Its Limits*, Oxford ; New York: Oxford University Press, 2008.
- Cohen, G A. *Rescuing Justice and Equality*, Cambridge, MA: Harvard University Press, 2009.



- Cook, John, Dana Nuccitelli, Sarah A Green, Mark Richardson, Bärbel Winkler, Rob Painting, Robert Way, Peter Jacobs, and Andrew Skuce. "Quantifying the Consensus on Anthropogenic Global Warming in the Scientific Literature." *Environmental Research Letters* 8, no. 2 (May 15, 2013): 024024.
- Copp, David. "Reasonable Acceptability and Democratic Legitimacy: Estlund's Qualified Acceptability Requirement." *Ethics* 121, no. 2 (January 2011): 239–269.
- Crouch, Colin. *Post-Democracy*, Cambridge: Wiley, 2004.
- de Tocqueville, Alexis. *Democracy in America*. Edited by Isaac Kramnick, Translated by Gerald Bevan, London: Penguin, 2003.
- Dewey, John. *The Public and Its Problems*, Athens, OH: Ohio University Press, 1927.
- Dewey, John. *Theory of Valuation*. International Encyclopedia of Unified Science, Vol. 2, Chicago, IL: Chicago University Press, 1939.
- Dewey, John, and James Hayden Tufts. *Ethics*. 2nd ed., New York: Henry Holt and Company, 1932.
- Dietrich, Franz. "The Premises of Condorcet's Jury Theorem Are Not Simultaneously Justified." *Episteme* 5, no. 1 (2008): 56–73.
- Dietrich, Franz, and Kai Spiekermann. "Epistemic Democracy with Defensible Premises." *Economics and Philosophy* 29, no. 1 (April 3, 2013): 87–120.
- Dietrich, Franz, and Kai Spiekermann. *Epistemic Democracy with Defensible Premises*. Working Papers, London: London School of Economics and Political Science, 2010.
- Downs, Anthony. *An Economic Theory of Democracy*, New York: Harper, 1957.
- Dworkin, Ronald. "What Is Equality? Part 4: Political Equality." *University of San Francisco Law Review* 22, no. 1 (Autumn 1987): 1–30.
- Elster, Jon. *Securities Against Misrule*, Cambridge: Cambridge University Press, 2013.
- Elster, Jon. *Solomonic Judgements: Studies in the Limitation of Rationality*, Cambridge: Cambridge University Press, 1989.
- Elster, Jon. "The Market and the Forum." In *Foundations of Social Choice Theory*, edited by Jon Elster and Aanund Hylland, 103–132, Cambridge: Cambridge University Press, 1986.
- Estlund, David M. *Democratic Authority: a Philosophical Framework*, Princeton, NJ: Princeton University Press, 2008.
- Estlund, David M. "Jeremy Waldron on Law and Disagreement." *Philosophical Studies* 99,

no. 1 (2000): 111–128.

Farrelly, Colin. “Virtue Epistemology and the ‘Epistemic Fitness’ of Democracy.” *Political Studies Review* 10, no. 1 (January 20, 2012): 7–22.

Fuerstein, Michael. “Epistemic Democracy and the Social Character of Knowledge.” *Episteme* 5, no. 1 (2008): 74–93.

Galston, William A. *Liberal Pluralism: the Implications of Value Pluralism for Political Theory and Practice*, Cambridge: Cambridge University Press, 2002.

Gaus, Gerald. *Justificatory Liberalism : an Essay on Epistemology and Political Theory*, Oxford: Oxford University Press, 1996.

Gaus, Gerald. “On Seeking the Truth (Whatever That Is) Through Democracy: Estlund's Case for the Qualified Epistemic Claim.” *Ethics* 121, no. 2 (January 2011): 270–300.

Gaus, Gerald. *The Order of Public Reason : a Theory of Freedom and Morality in a Diverse and Bounded World*, New York: Cambridge University Press, 2011.

Gaus, Gerald F. “Liberal Neutrality: a Compelling and Radical Principle.” *Critical Review* (1994).

Gigerenzer, Gerd. *Rationality for Mortals: How People Cope with Uncertainty*. Evolution and Cognition Series, New York: Oxford University Press, 2008.

Gigerenzer, Gerd, and Peter M Todd. *Simple Heuristics That Make Us Smart*, New York: Oxford University Press, 1999.

Goodin, Robert E, and David M Estlund. “The Persuasiveness of Democratic Majorities.” *Politics Philosophy Economics* 3, no. 2 (2004): 131–142.

Goodin, Robert E, and Kai Spiekermann. “Epistemic Aspects of Representative Government.” *European Political Science Review* 4, no. 3 (November 23, 2011): 303–325.

Granovetter, Mark. “Threshold Models of Collective Behaviour.” *American Journal of Sociology* 83, no. 6 (May 1978): 1420–1443.

Gray, John. *Mill on Liberty: a Defence*. Second Edition. London: Routledge, 1996.

Gray, John. *Two Faces of Liberalism*, New York: The New Press, 2000.

Green, Donald P. “On Evidence-Based Political Science.” *Daedalus* 134, no. 3 (2005): 96–100.

Green, Jeffrey E. *The Eyes of the People : Democracy in an Age of Spectatorship*, Oxford ; New York: Oxford University Press, 2010.

- Grofman, Bernard, Guillermo Owen, and Scott L Feld. “Thirteen Theorems in Search of Truth.” *Theory and Decision* 15, no. 3 (1983): 261–278.
- Hamilton, Alexander, James Madison, and John Jay. *The Federalist: with Letters of Brutus*. Edited by Terence Ball, Cambridge: Cambridge University Press, 2003.
- Hardin, Russell. *How Do You Know? : the Economics of Ordinary Knowledge*, Princeton NJ: Princeton University Press, 2009.
- Harford, Tim. *Adapt: Why Success Always Starts with Failure*, London: Little, Brown, 2011.
- Hayek, Friedrich A von. “The Use of Knowledge in Society.” *The American Economic Review* 35, no. 4 (September 1945): 519–530.
- Haynes, Laura, Owain Service, Ben Goldacre, and David Torgerson. *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*, Cabinet Office Behavioural Insights Team, 2012.
- Hempel, Carl G. *Philosophy of Natural Science*, Englewood Cliffs, NJ,: Prentice-Hall, 1966.
- Hershenov, David B. “Two Epistemic Accounts of Democratic Legitimacy.” *Polity* 37, no. 2 (April 2005): 216–234.
- Hibbing, John R, and Elizabeth Theiss-Morse. *Stealth Democracy: Americans' Beliefs About How Government Should Work*, Cambridge: Cambridge University Press, 2002.
- Hirschman, Albert O. *The Rhetoric of Reaction*, Cambridge, MA: Harvard University Press, 1991.
- Hobson, J A. “The Re-Statement of Democracy.” *Contemporary Review* no. 81 (January 1, 1902): 262–272.
- Holtug, Nils. “Egalitarianism and the Levelling Down Objection.” *Analysis* 58, no. 2 (1998): 166–174.
- Honderich, Ted. *Conservatism*, London: Pluto Press, 2005.
- Hume, David. “Of the Original Contract.” In *Essays, Moral, Political and Literary*, edited by Eugene F Miller, Indianapolis: Liberty Fund, 1987.
- Ingham, Sean. “Disagreement and Epistemic Arguments for Democracy.” *Politics, Philosophy and Economics* (2013).
- James, William. *Pragmatism. Great Books in Philosophy*, Amherst, MA: Prometheus Books, 1907.
- Kahneman, Daniel. *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux, 2011.

- Kahneman, Daniel, and Gary Klein. "Conditions for Intuitive Expertise: a Failure to Disagree." *American Psychologist* 64, no. 6 (September 2009): 515–526.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky. *Judgment Under Uncertainty : Heuristics and Biases*, Cambridge: Cambridge University Press, 1982.
- Kelly, Jamie Terence. *Framing Democracy : a Behavioral Approach to Democratic Theory*, Princeton NJ: Princeton University Press, 2012.
- Kestenbaum, David, and Baldur Hedinnsson. "A New Mom, Bjork's Dad and the President of Iceland." NPR: Planet Money Podcast, April 15, 2011.
- Keynes, John Maynard. "The General Theory of Employment." *The Quarterly Journal of Economics* 51, no. 2 (February 1937): 209–223.
- Kitcher, Philip. "Division of Cognitive Labor." *The Journal of Philosophy* 87, no. 1 (1990): 5–22.
- Kitcher, Philip. *Preludes to Pragmatism*, Oxford: Oxford University Press, 2012.
- Kitcher, Philip. "Public Knowledge and the Difficulties of Democracy." *Social Research* 73, no. 4 (Winter 2006): 1–20.
- Kitcher, Philip. *Science in a Democratic Society*, Amherst, New York: Prometheus Books, 2011.
- Kitcher, Philip. *The Ethical Project*, Cambridge, Mass.: Harvard University Press, 2011.
- Knight, Jack, and James Johnson. "Democratic Experimentalism." presented at Midwestern Political Science Association Annual Meeting, 1–40, 2013.
- Knight, Jack, and James Johnson. *The Priority of Democracy : Political Consequences of Pragmatism*, New York, Princeton N.J.: Russell Sage Foundation; Princeton University Press, 2011.
- Kukathas, Chandran. *The Liberal Archipelago*, Oxford: Oxford University Press, 2007.
- Kuran, Timur. "Now Out of Never: the Element of Surprise in the East European Revolution of 1989." *World Politics* 44, no. 1 (October 1991): 7–48.
- Kuran, Timur, and Cass Sunstein. "Availability Cascades and Risk Regulation." *Stanford Law Review* 51, no. 4 (1999).
- Landemore, Hélène. "Deliberation, Cognitive Diversity, and Democratic Inclusiveness: an Epistemic Argument for the Random Selection of Representatives." *Synthese* (April 4, 2012).
- Landemore, Hélène. *Democratic Reason*, Princeton, NJ: Princeton University Press, 2012.

- Landemore, Hélène. “Why the Many Are Smarter Than the Few and Why It Matters.” *Journal of Public Deliberation* 8, no. 1 (2012).
- Landemore, Hélène, and Jon Elster. *Collective Wisdom: Principles and Mechanisms*, Cambridge: Cambridge University Press, 2012.
- Lazarsfeld, Paul F. “The American Soldier-an Expository Review.” *The Public Opinion Quarterly* 13, no. 3 (Autumn 1949): 377–404.
- Levi, Isaac. *Hard Choices : Decision Making Under Unresolved Conflict*, Cambridge: Cambridge University Press, 1986.
- Lindblom, Charles E. *The Intelligence of Democracy: Decision Making Through Mutual Adjustment*. Vol. 363, New York: Free Press, 1965.
- Lindblom, Charles E. “The Science of ‘Muddling Through’.” *Public Administration Review* (1959): 79–88.
- Lippert-Rasmussen, Kasper. “Estlund on Epistocracy: a Critique.” *Res Publica* 18, no. 3 (August 2012): 241–258.
- List, Christian, and Robert E Goodin. “Epistemic Democracy: Generalizing the Condorcet Jury Theorem.” *The Journal of Political Philosophy* 9, no. 3 (2001): 277–306.
- List, Christian, and Philip Pettit. *Group Agency : the Possibility, Design, and Status of Corporate Agents*, Oxford: Oxford University Press, 2011.
- MacGilvray, Eric. “Experience as Experiment: Some Consequences of Pragmatism for Political Theory.” *American Journal of Political Science* 43, no. 2 (April 1999): 542–565.
- MacGilvray, Eric. *Reconstructing Public Reason*, Cambridge, MA: Harvard University Press, 2004.
- Mackie, Gerry. *Democracy Defended*, Cambridge: Cambridge University Press, 2003.
- Mackie, Gerry. “Rational Ignorance and Beyond.” In *Collective Wisdom: Principles and Mechanisms*, 1–44, Cambridge: Cambridge University Press, 2012.
- Madingley, Lord Browne of. *Securing a Sustainable Future for Higher Education: an Independent Review of Higher Education Funding & Student Finance*, 2010.
- Manin, Bernard. *The Principles of Representative Government. Themes in the Social Sciences*, Cambridge: Cambridge University Press, 1997.
- Mansbridge, Jane. “Rethinking Representation.” *American Political Science Review* 97, no. 4 (November 2003).
- Menasse, Robert. “Joseph Und Angela: Menasse Über Die EU.” *Kleine Zeitung*, March 27,

2010.

- Merton, Robert K. "The Unanticipated Consequences of Purposive Social Action." *American Sociological Review* 1, no. 6 (December 1936): 894–904.
- Mill, James, and Thomas Babington Macaulay. *Utilitarian Logic and Politics: James Mill's "Essay on Government," Macaulay's Critique, and the Ensuing Debate*. Edited by Jack Lively and John Rees, Oxford: Clarendon Press, 1978.
- Mill, John Stuart. "Considerations on Representative Government." In *Utilitarianism; on Liberty; Considerations on Representative Government; Remarks on Bentham's Philosophy*, edited by Geraint Williams, 187–428, London: Dent, 1993.
- Mill, John Stuart. "On Liberty." In *Utilitarianism; on Liberty; Considerations on Representative Government; Remarks on Bentham's Philosophy*, edited by Geraint Williams, London: J.M. Dent, 1993.
- Misak, Cheryl J. "A Culture of Justification: the Pragmatist's Epistemic Argument for Democracy." *Episteme* 5, no. 1 (2008): 94–105.
- Misak, Cheryl J. *The American Pragmatists*, Oxford: Oxford University Press, 2013.
- Misak, Cheryl J. *Truth, Politics, Morality: Pragmatism and Deliberation*. Vol. 31, Routledge, 2000.
- Mitchell, Melanie. *Complexity : a Guided Tour*, Oxford England ; New York: Oxford University Press, 2009.
- Mitchell, Sandra D. *Unsimple Truths : Science, Complexity, and Policy*, Chicago: University of Chicago Press, 2009.
- Mitchell, Sandra D, and Wolfgang Streeck. *Complex, Historical, Self-Reflexive: Expect the Unexpected! MPIfG Working Paper*, Cologne: Max-Planck-Institute for the Study of Societies, 2009.
- Morozov, Evgeny. *To Save Everything, Click Here: the Folly of Technological Solutionism*, New York: PublicAffairs, 2013.
- Morton, Rebecca B, and Kenneth C Williams. "Experimentation in Political Science." In *The Oxford Handbook of Political Methodology*, edited by Janet Box-Steffensmeier, David Collier, and Henry Brady, 1–27, Oxford: Oxford University Press, 2006.
- Nozick, Robert. *Philosophical Explanations*, Cambridge, MA: Harvard University Press, 1981.
- Ober, Josiah. *Democracy and Knowledge: Innovation and Learning in Classical Athens*, Princeton, NJ: Princeton University Press, 2008.

- Ober, Josiah. "Democracy's Wisdom: an Aristotelian Middle Way for Collective Judgment." *American Political Science Review* (n.d.).
- Page, Benjamin I, and Robert Y Shapiro. *The Rational Public : Fifty Years of Trends in Americans' Policy Preferences. American Politics and Political Economy Series*, Chicago: University of Chicago Press, 1992.
- Page, Scott E. *Diversity and Complexity. Primers in Complex Systems. Primers in Complex Systems*, Princeton, NJ: Princeton University Press, 2011.
- Page, Scott E. *The Difference*, Princeton, NJ: Princeton University Press, 2007.
- Page, Scott E. "Uncertainty, Difficulty, and Complexity." *Journal of Theoretical Politics* 20, no. 2 (March 2008): 115–149.
- Peter, Fabienne. *Democratic Legitimacy. Routledge Studies in Social and Political Thought*, New York: Routledge, 2009.
- Peter, Fabienne. "Pure Epistemic Proceduralism." *Episteme* 5, no. 1 (2008): 33–55.
- Pettit, Philip. "Depoliticizing Democracy." *Ratio Juris* 17, no. 1 (March 2004): 52–65.
- Pettit, Philip. *On the People's Terms*, Cambridge: Cambridge University Press, 2012.
- Pettit, Philip. "The Common Good." In *Justice and Democracy: Essays for Brian Barry*, Cambridge: Cambridge University Press, 2004.
- Posner, Eric A, and Adrian Vermeule. *The Executive Unbound: After the Madisonian Republic*, Oxford: Oxford University Press, 2011.
- Putnam, Hilary. *The Collapse of the Fact/Value Dichotomy: and Other Essays*, Cambridge, MA: Harvard University Press, 2002.
- Rawls, John. *Political Liberalism*, New York: Columbia University Press, 2005.
- Rehfeld, Andrew. "The Child as Democratic Citizen." *The Annals of the American Academy of Political and Social Science* 633, no. 1 (January 2011): 141–166.
- Richardson, Henry S. "Truth and Ends in Dewey's Pragmatism." *Canadian Journal of Philosophy. Supplementary Volume* 24 (1998): 109–147.
- Riedel, Stefan. "Edward Jenner and the History of Smallpox and Vaccination." *Baylor University Medical Center Proceedings* 18, no. 1 (2005): 21–25.
- Riker, William H. *Liberalism Against Populism : a Confrontation Between the Theory of Democracy and the Theory of Social Choice*, Long Grove, IL: Waveland Press, 1982.
- Rittel, Horst W J, and Melvin M Webber. "Dilemmas in a General Theory of Planning."

*Policy Sciences* 4, no. 2 (June 1973): 155–169.

Robin, Corey. *The Reactionary Mind: Conservatism From Edmund Burke to Sarah Palin*, Oxford: Oxford University Press, 2011.

Rodrik, Dani. “The New Development Economics: We Shall Experiment, but How Shall We Learn?.” *HKS Faculty Research Working Paper Series* (2008).

Roth, Amanda. “Ethical Progress as Problem-Resolving.” *Journal of Political Philosophy* 20, no. 4 (May 2, 2012): 384–406.

Sabel, Charles F, and William H Simon. “Minimalism and Experimentalism in the Administrative State.” *Georgetown Law Journal* 100 (2011): 53–93.

Sabel, Charles F, and Jonathan Zeitlin. “Experimentalism in the EU: Common Ground and Persistent Differences.” *Regulation & Governance* 6, no. 3 (September 20, 2012): 410–426.

Sabel, Charles F, and Jonathan Zeitlin. “Experimentalist Governance.” In *The Oxford Handbook of Governance*, edited by David Levi-Faur, Oxford University Press, 2011.

Sagalnik, Matthew J, Peter Sheridan Dodds, and Duncan J Watts. “Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market.” *Science* 311 (February 2006): 854–856.

Sagar, Rahul. *Secrets and Leaks: the Dilemma of State Secrecy*, Princeton, NJ: Princeton University Press, 2013.

Saunders, Ben. “Combining Lotteries and Voting.” *Politics, Philosophy and Economics* 11, no. 4 (October 31, 2012): 347–351.

Saunders, Ben. “Democracy, Political Equality, and Majority Rule.” *Ethics* 121 (October 2010): 148–177.

Schelling, Thomas C. *The Strategy of Conflict*, Cambridge, MA: Harvard University Press, 1980.

Schumpeter, Joseph A. *Capitalism, Socialism and Democracy*, London: Routledge, 2012.

Sen, Amartya. “Democracy as a Universal Value.” *Journal of Democracy* 10, no. 3 (1999): 3–17.

Sen, Amartya. *The Idea of Justice*, Cambridge, MA: Harvard University Press, 2009.

Silver, Nate. *The Signal and the Noise: The Art and Science of Prediction*, (London: Allen Lane, 2012).

Somin, Ilya. “Voter Ignorance and the Democratic Ideal.” *Critical Review* 12, no. 4



(Autumn 1998): 413–458.

- Spufford, Francis. *Red Plenty*, London: Graywolf Press, 2012.
- Sunstein, Cass R. *Going to Extremes : How Like Minds Unite and Divide*, Oxford: Oxford University Press, 2009.
- Sunstein, Cass R. *Infotopia: How Many Minds Produce Knowledge: How Many Minds Produce Knowledge*, Oxford: Oxford University Press, USA, 2006.
- Surowiecki, James. *The Wisdom of Crowds*. 1st ed., New York: Doubleday, 2004.
- Taleb, Nassim Nicholas. *Antifragile: Things That Gain From Disorder*, New York: Random House, 2012.
- Taleb, Nassim Nicholas. *The Black Swan : the Impact of the Highly Improbable*. 1st ed., New York: Random House, 2007.
- Talisso, Robert B. *A Pragmatist Philosophy of Democracy. Routledge Studies in Contemporary Philosophy*, New York, London: Routledge, 2007.
- Talisso, Robert B. *Pluralism and Liberal Politics*, New York: Routledge, 2012.
- Tetlock, Philip E. *Expert Political Judgment: How Good Is It? How Can We Know?*, Princeton, NJ: Princeton University Press, 2005.
- Tuck, Richard. *Free Riding*, Cambridge, MA: Harvard University Press, 2009.
- Urbinati, Nadia. *Representative Democracy : Principles and Genealogy*, Chicago, IL: University of Chicago Press, 2006.
- Urbinati, Nadia, and Maria Paula Saffon. “Procedural Democracy, the Bulwark of Equal Liberty.” *Political Theory* (February 26, 2013).
- Valentini, Laura. “Justice, Disagreement and Democracy.” *British Journal of Political Science* 43, no. 1 (July 19, 2012): 177–199.
- Van Parijs, Philippe. “The Disfranchisement of the Elderly, and Other Attempts to Secure Intergenerational Justice.” *Philosophy and Public Affairs* 27, no. 4 (Autumn 1998): 292–333.
- Vermeule, Adrian. “Collective Wisdom and Institutional Design.” In *Collective Wisdom and Institutional Design*, edited by Jon Elster and Hélène Landemore, 338–368, Cambridge: Cambridge University Press, 2012.
- Vermeule, Adrian. *Mechanisms of Democracy: Institutional Design Writ Small*, Oxford: Oxford University Press, 2007.

- Vibert, Frank. *The Rise of the Unelected: Democracy and the New Separation of Powers*, Cambridge: Cambridge University Press, 2007.
- Viehoff, Daniel. "Debate: Procedure and Outcome in the Justification of Authority." *Journal of Political Philosophy* 19, no. 2 (June 2011): 248–259.
- Waldron, Jeremy. *Law and Disagreement*, Oxford: Oxford University Press, 2001.
- Wall, Steven. "On Justificatory Liberalism." *Politics, Philosophy and Economics* 9, no. 2 (April 28, 2010): 123–149.
- Wall, Steven. "Public Reason and Moral Authoritarianism." *The Philosophical Quarterly* 63, no. 250 (December 17, 2012): 160–169.
- Watts, Duncan J. *Everything Is Obvious: Once You Know the Answer*. 1st ed., New York: Crown Business, 2011.
- Wiens, David. "Prescribing Institutions Without Ideal Theory." *Journal of Political Philosophy* 20, no. 1 (2012): 45–70.