

THE PSYCHOLOGY OF DEMOCRATIC DELIBERATION: FROM PRACTICE TO SYSTEM

by

David Moscrop

B.Soc.Sci (Honours), University of Ottawa, 2007

M.A., University of Ottawa, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES AND POSTDOCTORAL STUDIES

(Political Science)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2017

©David Moscrop, 2017

## **Abstract**

Accounts of democratic deliberation assume and require citizens who are capable of rational and autonomous cognition. Such individuals are expected to be able to gather, process, and communicate information in such a way that allows them to accurately account for their preferences, including providing reasons for those preferences. The epistemic defense of deliberative democracy suggests that this is possible and that citizens who deliberate can generate good judgments and decisions.

In this dissertation, I bring findings from social and political psychology to bear on the question of whether citizens can make good judgments and decisions through democratic deliberation. Data collected over the last five decades casts some doubts over whether they can. However, as I argue, there is good reason to believe that deliberation, despite these challenges, is often superior to alternative approaches to decision making and that, moreover, there are individual practice and institutional design responses that can mitigate the deleterious effects of phenomena that bring about cognitive distortion, bias, and error when citizens deliberate.

In the first section of this dissertation, I argue that the epistemic defense of deliberation—including the need for rational, autonomous citizens—is challenged by findings from social and political psychology, but that democratic deliberation remains a possible and superior form of public judgment and decision making. In the second section, I use institutional theory, deliberative systems literature, and findings from psychology to discuss ways of thinking about autonomy and deliberation, and I develop approaches to limiting or overcoming the

challenges mentioned in section one. These approaches are rooted in both broader institutional design and deliberative system design and in specific deliberative practices.

## **Preface**

This dissertation is original, unpublished work by the author, David Moscrop.

# Table of Contents

<b>Abstract .....</b>	<b>ii</b>
<b>Preface .....</b>	<b>iv</b>
<b>Table of Contents .....</b>	<b>v</b>
<b>List of Figures .....</b>	<b>x</b>
<b>Acknowledgements .....</b>	<b>xi</b>
<b>Dedication .....</b>	<b>xii</b>
<b>Chapter 1: Can we deliberate? The problem of rationality and autonomy in democratic societies .....</b>	<b>1</b>
<b>Judgments and decisions.....</b>	<b>4</b>
<b>Reason giving, good/correct judgments and decisions, and deliberative democracy.....</b>	<b>6</b>
<b>Rationality and autonomy.....</b>	<b>10</b>
<b>Epistemic democracy .....</b>	<b>15</b>
<b>Nature, culture, and deliberation .....</b>	<b>17</b>
<b>Affect and autonomy in political philosophy.....</b>	<b>20</b>
<b>Cognitive limitations, distortions, and bias in political contexts .....</b>	<b>27</b>
<b>Models from social and political psychology .....</b>	<b>33</b>
<b>Chapter outline .....</b>	<b>36</b>
<b>Chapter 2: The Epistemic argument for democratic deliberation and the need for autonomy.....</b>	<b>46</b>
<b>A note on truth.....</b>	<b>51</b>
<b>Deliberative democracy as epistemology.....</b>	<b>54</b>
<b>Politics and expertise in deliberations .....</b>	<b>57</b>
<b>Two more reasons why deliberation is a more rational way to generate judgments and decisions.....</b>	<b>60</b>
<b>Epistemic proceduralism: the best of two worlds? .....</b>	<b>63</b>

Elstub’s critique of the epistemic defense of deliberation .....	67
Autonomy in deliberative democracy: the keystone capacity .....	69
Autonomy and the principle of non-self-deception .....	77
Self-determination and democracy .....	83
Relational autonomy .....	86
A note on the philosophical debate between free will and determinism .....	88
Conclusion and summary .....	90
<b>Chapter 3: Challenges to autonomy in democratic deliberation .....</b>	<b>91</b>
A note on epistemology, the brain, and our environment .....	94
What is expected and required of deliberative agents? .....	96
The expression of rationality and self-determination/autonomy .....	97
Self-determination and autonomy .....	100
Validity of reasons .....	102
More legitimate outcomes .....	104
From evolutionary and social psychology to deliberation: dual process cognition in the system 1 and system 2 modes .....	105
The systems approach: how we generate judgments and beliefs .....	110
What is cognitive distortion and bias? .....	112
Models of cognition .....	116
<i>The elaboration likelihood model (ELM): persuasion and opinion/attitude change</i> .....	116
<i>System justification</i> .....	120
<i>Social intuitionism</i> .....	128
<i>Automaticity</i> .....	137
Now what? Options for responding to challenges to autonomous deliberation .....	145
Conclusion and Summary .....	146
<b>Chapter 4: Motivated reasoning and democratic deliberation .....</b>	<b>149</b>

<b>Motivated reasoning</b> .....	151
<i>Overview</i> .....	151
<i>Hot cognition</i> .....	155
<i>Directional goals, accuracy goals, and reason-giving</i> .....	156
<b>The model of motivated reasoning in summary</b> .....	166
<b>Cognitive models, motivated reasoning, and democratic deliberation</b> .....	169
<b>Conclusion and Summary</b> .....	175
<b>Chapter 5: Institutional deliberative autonomy and practice-based responses to the psychological challenges of democratic deliberation</b> .....	176
<b>How institutional design can be used to respond to challenges to autonomy and</b>	
<b>rationality</b> .....	177
<b>Tradeoffs and limits: Who can and who cannot deliberate</b> .....	180
<b>Habermas and autonomy</b> .....	183
<b>Institutional deliberative autonomy</b> .....	187
<i>The problem</i> .....	187
<i>The solution</i> .....	191
<b>How to evaluate IDA</b> .....	197
<i>Assessment one: Discourse Quality Index</i> .....	201
<i>Assessment two: expert qualitative reports</i> .....	203
<i>Assessment three: Self-reporting</i> .....	205
<b>Summary of institutional deliberative autonomy</b> .....	207
<b>Concepts for deliberative design within the context of institutional deliberative</b>	
<b>autonomy</b> .....	209
<i>A note on facilitation</i> .....	209
<i>Iteration</i> .....	211
<i>A-rational receptivity</i> .....	215
<i>Cognitive diversity</i> .....	218
<i>Targeted motivation</i> .....	221
<b>Next steps for deliberative institutional design: hypotheses and tests</b> .....	223
<b>Conclusion and Summary</b> .....	224
<b>Chapter 6: Democratic deliberation and institutions</b> .....	226

<b>What is an institution and why do institutions matter?</b> .....	227
<i>The rules of the game: Denzau and North</i> .....	228
<i>The ontology of institutions: Searle</i> .....	230
<b>How do institutions and cognition interact?</b> .....	234
<i>Ontological security: What is there in the world?</i> .....	236
<i>Meaning maintenance: Start making sense?</i> .....	238
<b>How do institutions encourage poor reasoning, judgments, and decisions?</b> .....	242
<i>Campaigns and elections</i> .....	243
<i>The media and attitude formation</i> .....	249
<i>Political parties and partisanship</i> .....	251
<i>Electoral systems</i> .....	255
<i>Ideology and hegemony: a special case</i> .....	258
<b>Conclusion and summary</b> .....	264
<b>Chapter 7: A deliberative systems approach to generating epistemically better democratic deliberation</b> .....	266
<b>Understanding the deliberative system: Concerns, functions, and roles</b> .....	267
<i>What are the concerns of a deliberative system?</i> .....	267
<i>What are the functions of a deliberative system?</i> .....	268
<i>What are the roles of experts, policy-makers, and citizens in a deliberative system?</i> .....	272
<b>Deliberative systems and cognitive error</b> .....	276
<b>More deliberations, integrated deliberations</b> .....	278
<b>Situated and balanced technocracy</b> .....	282
<b>Stability and disruption</b> .....	287
<b>Experiments in deliberation</b> .....	289
<b>Conclusion and summary</b> .....	291
<b>Chapter 8: Practicing democratic deliberation</b> .....	293
<b>Why bother with epistemically good democratic deliberation?</b> .....	294
<b>Implications and further research</b> .....	295
<b>Next steps</b> .....	299
<b>The place of deliberation in our political lives</b> .....	301
<b>Deliberation, decency, and human capacity</b> .....	303

**Bibliography.....304**

## List of Figures

Figure 1. <i>Institutional deliberative autonomy and its components</i> .....	195
---	-----

## Acknowledgements

Writing a dissertation is at times a lonely endeavour, but it is nonetheless a collective one. In some ways, writing this has been the most enjoyable part of my graduate school experience—of the official bits of my program, anyway. I owe this fact to several people—far more than I can name here. However, a few stand out. First, I wish to acknowledge and thank my supervisor, Mark E. Warren. I came to the University of British Columbia mostly because of him, and I'm glad I did. His help in the process has been invaluable. Mark has a singularly remarkable way of synthesizing, simplifying, and communicating complex ideas. I've been the beneficiary of these skills, and I'm awfully thankful for it.

Next, I thank my committee members. Andrew Owen and Steven Heine have been at once rigorous and encouraging in supporting a dissertation that includes research from philosophy, neuroscience, political theory, social and political psychology, and other fields. Their patience and insights have, along with Mark's, made this dissertation much, much better than I could have made it on my own. I chose my committee to keep me honest; I figured that you could perhaps fool one expert in one field about research from another area of study—but it would be hard to fool three remarkably talented experts from three different areas of study. I chose...wisely.

Finally, I thank my family and friends; there far too many folks to name here—other than, of course, my mom, whose support has been instrumental in me making it this far in the first place. They know who they are. My deepest thanks to them.

## **Dedication**

To the family I have lost during these years of study. I was so lucky to have you with me at the beginning; I wish you were here at the end.

# Chapter 1: Can we deliberate? The problem of rationality and autonomy in democratic societies

Can individuals make rational, autonomous political judgments in democratic deliberative settings? That is the question I ask and answer in this dissertation. To ask this question is to ask two related, but independent questions. First: Are there available, sufficient, and meaningful opportunities for citizens to become educated about political issues; to come together in a setting free from coercion—emotional, psychological, or physical—to discuss their preferences, desires, and goals; and to give reasons for and against propositions in an attempt to reach considered, just, and legitimate decisions? And second: Are human beings equipped with or able to cultivate the necessary cognitive capacities for rationally and autonomously engaging in democratic deliberation with one another? While the first question raises significant ethical and practical concerns about democratic systems and the role of individual citizens in collective self-determination and self-government, it is the second question that I am interested in here. Why? Because the second question points to concerns about the extent to which human beings are capable of reaching the rational and autonomous political judgments and decisions that proponents of deliberative democracy argue that they can, and will, reach when deliberating. So, the viability of theories of deliberative democracy thus depend, in part, on the answer to this second question.

Concerns about the capacity of deliberators to make good judgments and decisions emerge because there is a gap between what we normatively expect or desire from those individuals—critical, factual, and relevant judgments and decisions whose origins are rooted *primarily* in the considered and expressed considerations of the agent, which are themselves

drawn from *known* internal or external motivations—and what they tend to be able to deliver day-to-day. This dissertation is about what creates that gap and how we might go about bridging it. Given this concern with autonomy in democratic deliberation, in this dissertation I spend lots of time discussing human cognition—its individual and social manifestations—and how it relates to both the normative (i.e. desired) and practical (i.e. necessary) requirements of deliberative citizenship. In the following pages, while I address both the capacities of autonomy and rationality I pay particularly close attention to autonomy and argue that it is the keystone capacity for good deliberation. Autonomy is central to deliberation because while both it and rationality are important deliberative capacities, *autonomy is necessary to deliberation in a special way*. Democratic deliberation requires that individuals put reasons on the table for discussion and debate. While an individual might have irrational reasons or justifications for their preferences (which are not ideal from a deliberative point of view), if those irrational considerations can be put on the table, they can at least be interrogated by the group. The ability to deliberate in this way requires, as I will argue below and more extensively in chapter two, a capacity for autonomy. So, autonomy allows rational/irrational reasons to be put on the public register and taken up by those who deliberate. That is why I argue it is a keystone capacity, and why I spend extra time analyzing it vis-à-vis deliberation.

Again, the primary argument I make in this dissertation is: *When it comes to democratic deliberation, there is a gulf between what we expect and required of citizens, and what our cognitive capacities and the institutions that they shape/that shape them can deliver*. In short: we, as individuals who deliberate in democratic settings, are expected to make better judgments and decisions than we *tend to make* outside of them, on our own or with others. We

are supposed to be thoughtful and critical; we are expected to adopt a higher standard of critical thinking than if we were just popping into the grocery store to choose some vegetables and meat for dinner, or picking which film to watch for an evening, or choosing between the red shirt and the blue one. But often we fail to live up to those higher standards. Often, something gets in the way of us making good judgments and decisions. Now, this argument could —no surprise—also apply to all kinds of human behaviour if there were good reason to apply it, including those just mentioned: how we eat, how we purchase goods and services, what we watch, or even how we choose a partner. But in this dissertation, I am specifically concerned with how certain institutional and cognitive structures and phenomena *interact with individuals as potential participants in public, political deliberation as outlined by contemporary deliberative democracy theory*. Proponents of democratic deliberation argue that deliberation is a good way to generate political outcomes; I want to put that claim to the test in light of data from political and social psychology that (often implicitly, when it comes to deliberation) suggests otherwise.

My focus in this dissertation is thus on the relationship between individuals, institutions, and deliberative democracy vis-à-vis the *epistemic function* deliberation serves as a democratic practice. And while, as I argue, there are reasons to be sceptical about the extent to which individuals tend to make “good” judgments and decisions—a term I define below—there is also reason to be optimistic about improving those judgments and decisions through careful personal practice and institutional reform. To get to that point, I first call into question the extent to which individuals can engage in epistemically good democratic deliberation. I explore individual cognitive tendencies—especially those that distort or bias our thinking—drawing

from empirical evidence found in social and political psychology. I explore how these tendencies interact with the social and political institutions found in contemporary democratic societies and argue that to maximize the quality of judgments and decisions made in deliberative settings, we require changes in our personal practice, institutional design, and deliberative systems. Nonetheless, despite significant challenges to rationality and autonomy posed by certain cognitive tendencies, I argue that with good practice and design, we can achieve and maintain a high standard of epistemically valuable democratic deliberation.

## **Judgments and decisions**

In this dissertation, when I discuss a “judgment,” I mean to refer to a *single, considered conclusion reached by an individual*. For instance, I might reach the judgment that raising taxes on the richest two percent is good policy since we can leverage the relatively minimal marginal cost to those individuals for proportionately greater goods for social programs to help poorer people. That is a considered judgment. It is personal, discrete, and internal. Related to a judgment, but distinct from it, is a “decision”—by which I mean a *deliberate, collective choice to undertake (or to reject) some action or series of actions*. So, a legislative body might take up the judgment(s) of one or more individuals on, say, this matter of raising taxes and indeed make changes to legislation that does just that. Decisions, as I mean them in this dissertation, are collective and external, though they are enabled by judgments.

Individual judgments are separate and distinct from collective decisions, even though they likely contribute to bringing about a decision. In a debate or deliberation, individual judgments will affect group outcomes, though they may be of different sorts: moral, technical,

strategic, and so forth. So, for instance, a group might be made up of individuals who make different and incompatible judgments, but who take a collective decision for any number of reasons. As an example, imagine a deliberation around what to do about climate change. Jennifer reaches the judgment that a carbon tax is needed to address climate change; William prefers a cap-and-trade system. Each explains their judgment—their preference for one system over the other—and gives reasons for it. The body of which they are a part decides to choose cap-and-trade. While Jennifer and William reached separate judgments, the body made a single decision. If the body is required to vote, perhaps William and Jennifer will make a deal in which Jennifer supports William, which is a different sort of judgment leading to a different decision (e.g. to compromise or bargain in exchange for some other good). The examples and variations of examples of this distinction are endless, but the important takeaway is that judgments and decisions are distinct, but related in complex and variable ways.

Typically, however, and especially when it comes to democratic deliberation, a judgment is in part the result of an epistemic process of collecting, interpreting, and sharing information, and then reaching a considered conclusion that made lead to a decision—which, again, for the purposes of this dissertation, is a shared, single act of a deliberative or other decision-making body. This distinction is important to remember because throughout this dissertation I will argue that by improving individual judgments—that is, by improving the epistemic value of deliberation by enhancing individual capacities for rationality and autonomy—we can improve collective decision making. This argument rests on the claim that if we can improve the quality of information, preferences, types of reasoning and the sorts of reasons that emerge from those combined considerations, we can better facilitate shared

understanding, mutual respect, legitimacy, and the representativeness of decisions. (Note that the qualities of a good decision include non-cognitive goods.)

Now, to the heart of the matter: to answer the question of whether our cognitive capacities of rationality and autonomy are sufficient for the type of deliberative democracy advocated by leading theorists of democracy, I ask the following research question: *What is the impact of a-rational cognitive processes—cognitive processes that precede or circumvent autonomous, rational reflection, but which may have an impact on conscious thought—on the possibility of producing epistemically good judgments and decisions in deliberative settings?* I also ask, in response to challenges to autonomous and rational deliberation posed by such processes: *How can political theory and political science respond to the effects of such phenomena in ways that underwrite capacities of autonomy and rationality in relation to deliberation?* By now you might think that I assume the answer to the question “Can individuals make rational, autonomous, political judgments in democratic deliberative settings” is “No.” Close, but that is not quite true. I believe the answer is “Not always and not particularly well; but there are ways we can do better.” Indeed, the latter half of this dissertation will be specifically about how we can do better through changes to our personal practices and our institutions.

### **Reason giving, good/correct judgments and decisions, and deliberative democracy**

At this point, a bit more definitional clarity and precision are required. What is a *rational, autonomous judgment or decision*? And what is an *epistemically “good” judgment or decision*?

Answering each of these questions is essential since a great deal depends upon how one understands each of these concepts; accordingly, I will carefully address them in detail below. First, though, it is important to broadly define what it is I mean by “deliberative democracy”—since it is within the context of this family of theories of democracy that I will ask these questions—and why deliberation is critical to democratic theory in the context of contemporary liberal democracies (and elsewhere for that matter). While deliberative democracy is a contested concept, one of its core elements is common among the many theories of deliberation—enough that we can say it is accepted and held in common: reason giving.

Reason giving is a central requirement to deliberative democracy; indeed, the practice is at the very heart of deliberation as-such (Bohman 1998; Gutmann and Thompson 2004; Schneiderhan and Khan 2008; Warren 2002). Reason giving refers to the exchange of reasons among participants in a deliberation for or against a claim or proposal such that each is accountable to the others concerning *why they want what they want*. Reasons are given for the purposes of justification, but also so that individual preferences might be *produce, transformed, and/or justified to others*. The deliberative reason-giving approach stands in contrast to the mere aggregation of preferences (e.g. by a tally), which only requires that a participant in a political exercise states *what they want* (typically before a vote is taken to decide the matter).

The concept and practice of reason giving are important for this dissertation. As I will discuss in more detail in chapters three and four, reasons are essential to deliberation because they are the means by which individuals in a deliberation communicate. Indeed, in whatever

way they are delivered, which may be in a better or worse way depending on the individual and how a given deliberation is structured and carried out, reasons enable deliberation. Reasons allow for “deontic scorekeeping”: essentially a tally kept of what people say to one another, which each takes as binding the other to their utterance and its implications (Brandom 1994). Reasons also generate democratic goods including trust, coordinated actions, motivation to engage in future exchanges, and compliance. Reasons, when exchanged honestly and in a constructive way, create the foundations on which democratic deliberation rests—or, if you prefer, serve as a currency exchanged between participants. But, just as importantly, as I will argue throughout this dissertation, when we connect reasons to autonomous judgments and decisions, participants in deliberation become more likely to produce better outcomes.

What kind of better outcomes? Warren (2002) argues that democratic deliberation—*through “the give and take of reasons”*—is concerned with bringing about three types of outcomes: political, ethical, and epistemic. First, deliberative democracy’s *political* aims are related to generating better governance institutions and forming citizen preferences with an emphasis on reaching acceptable decisions based on reason giving; this approach stands in contrast to alternative (possibly complementary) approaches to generating decisions, including aggregating existing preferences through voting or bargaining, or by coercion. Next, deliberative democracy’s *ethical* function and related outcomes are bound up in the deontological claim that individuals ought to be treated as ends rather than means, and thus should be a part of the decision-making process when it comes to decisions that will affect them—whether the questions at hand are political or moral. This ethical commitment is part of the “all affected principle,” by which all of those affected by a decision are accepted as having a

moral claim to be a part of the process by which that decision is made. Finally, deliberation's *epistemic* function is based on the argument that deliberative democracy provides a more rational and thus better way of making decisions than expert/elite, power-based, or strategic approaches to decision making, such as technocracy, coercion or preference-aggregation. This claim is grounded in the argument that deliberation is better at producing collective judgments based on shared knowledge and understanding (hence the "better" outcomes). Knowledge generated in this way also tends to enjoy the benefit of being valid (what it takes for a statement to be considered valid in the context of deliberation is something I will examine below).

In this dissertation, this epistemic function of deliberative democracy is most important to me since I am interested how good or correct judgments and decisions can be brought about by rational, autonomous deliberators—which requires that deliberations generate epistemically reliable information flows and uptake. Accordingly, I will focus primarily on the epistemic dimension of deliberation and the epistemic goods that emerge from good deliberation. While the three functions of deliberative democracy may be related and mutually reinforcing, I do not spend much time exploring the ways they are related, except for a few interludes where warranted (and explicitly noted).

It is within this context of the epistemic dimensions of democratic deliberation that I define a "good" judgment and a "good" decision. A "good" *judgment* in deliberation is transparent, valid, and reliable. It is transparent when the motivations of the individual making it are *accessible and known to her and, if required, to others*; it is *valid* such that the premises and/or reasons upon which it is based logically lead to the conclusion in a way that can be

understood and accepted by others; and it is *reliable* in such a way that one should reasonably expect an individual to reaffirm it in the future given the same facts and context. A good *decision* is one that is taken collectively based on the compatible principles of good individual judgments, and which is collectively scaled so that the outcome is based on, as much as possible, accepted epistemic foundations (even though individual judgments may vary).<sup>1</sup> A good decision should meet the criteria set for good judgment. If in translating many judgments into a decision something is lost, and the criteria cannot be met, the decision is not a good one.<sup>2</sup> It is important to note that I do not intend the usage of the concept of “good” in any way that relies on the correspondence theory of truth in which an outcome “matches” the world “as it is.”

## **Rationality and autonomy**

When it comes to reasoning—which is an important part, but not the only part, of how we generate preferences reasons to support them—I do not assume that the *method* of reasoning must be based on either ends-means rationality or utility maximization. By this I mean that *I do not consider instrumental rationality* to be the *only* sort of rationality that matters when making political judgments and decisions in a deliberative context. The key issue regarding deliberation

---

<sup>1</sup> The end goal of deliberation is not consensus, but mutual understanding and acceptance of whatever outcomes are reached by those who deliberate. The same focus applies here since we can base a good decision on many different, but nonetheless good, individual judgments. Any collective decision that deliberators reach may or may not be compatible with the *substance* of each judgment, but should be compatible with the principles of that judgment.

<sup>2</sup> There are other criteria for a good decision relevant to democratic norms and deliberative democracy goods (e.g. publicity, legitimacy, etc.). They are in some ways related to my concerns here, but I have bracketed any discussion of them since I am primarily concerned with the *epistemic* function of deliberation.

is whether there is a *defensible logic at work during an exchange* that reflects a) both a basic understanding of the factual and normative reality of the world and b) which can be shared, understood, and in principle accepted by others under reasonable circumstances. Within this approach, there is also such a thing as “more or less” good (i.e. better or worse) in a given context. As we will see in chapter five, there is a “sliding scale” when it comes to one’s capacities for rationality and autonomy, and there is also such a scale for judgments and decisions (within a given system). This is to say that, on balance, given different approaches to reasoning and different interpretations of “facts” about the natural world and about the normative world *in the context of politics*, there will be space for dispute about what counts as good; accordingly, there is no scientific test for, or objective measure, that signifies the “right” or “true” way of judging or deciding—especially when it comes to ethical or moral matters in a society marked by deep and persistent disagreement.<sup>3</sup> However, *within the context of a given and generally-accepted system*, some judgments and decisions will be: i) closer to reasonable interpretations about what motivates the agent who decides; ii) based on more or less defensible accounts about what exists in the world and how it exists alongside and interacts with other things and phenomena; iii) backed up by more or less honest and accurate accounts about the motivation(s) one has for holding the preferences that generate a judgment or decision; iv) more or less likely to be reproduced in the future given the same or very similar contexts; and v) more or less accessible to other agents who have a political and moral right to participate in collective decision making about the political and social world they share and

---

<sup>3</sup> Although science as a way of accumulating knowledge is characterized by procedures for establishing validity intersubjectively.

whose rights are better respected by being a part of a political practice to which they have access to a shared epistemic foundation for decision making.

Of course, there are contexts in which conceiving of a capacity for rationality in terms of ends-means or utility-maximization makes sense—especially if that happens to be how some individuals think in some circumstances or how groups elect to work through an issue or some set of issues. For the purposes of this dissertation I am most often, unless otherwise stated, specifically interested in how ordinary (i.e. non-elite) individuals cognitively navigate and make sense of the complex political world day-to-day; and I am especially interested in how they think, judge, and decide when they are asked to engage in democratic deliberation, however they may tend to (cognitively) do so. But cognition tends to rely on a human capacity for rationality (and autonomy) that is constrained or “bounded” (March and Simon 1958). Accordingly, my argument is grounded in the belief—and supported by the literature that I will explore throughout these pages—that rationality is inherently *bounded* in the sense of being constrained by real-world limitations that limit ends-means/utility-maximizing rationality. My argument is also based on the idea that rationality is also inherently and inextricably *bound up with a-rationality* to some extent. And while rationality may be context-dependent and shaped by structures of authority and power (Flyvbjerg 1998), there must be nonetheless a basic shared contextual rationality in specific contexts if we are to achieve stable and widely understood—though not necessarily permanent or universally-agreed-upon—democratic decisions, though navigating how we conceive of rationality in different contexts will require some flexibility and openness to adaptation.

Accordingly, the conception of rationality that I rely on for this dissertation is broad; by rationality, I refer to the human capacity to make stable sense of the world *alongside others with whom they live in a social and political community*. This conception of rationality requires that individuals be able to draw *facts* from the world in a more-or-less consistent, reliable, and objectively accurate way (within a given system of meaning), and that they be able to communicate them to others on basic shared grounds (determined epistemically); by implication, this conception of rationality also involves that individuals are more or less able to agree on *some* normative facts—which can be decidedly trickier in a political context. My conception of rationality borrows a bit from the instrumental sense of the term mentioned above, but is balanced by a need for also including a communicative rationality perspective, which involves “processes of discussion and persuasion” to “[help] form bonds of understanding” between individuals with different histories, beliefs, and political and scientific understandings of the world (Parkinson 2006: 127; see also Dryzek 1990; Habermas 1996). Rationality, for my purposes, is thus individually *held* as a capacity, but publicly *shaped and shared* through interpersonal communication. Rationality in this context is a personal capacity but it is other-dependent. Why? Because as I conceive of it for the purposes of this dissertation, rationality depends on an intersubjective process of validation—we come to know that a judgment is rational through exchange with others in a given system, which is essential to democratic (and other sorts of) life.

Another, related, approach to conceiving of rationality is practice-based and aimed at navigating the world through *practical* understanding. This conception is often known as the “practical wisdom” approach to rationality. Grounded in Aristotle (1999 [350 B.C.E.]) and his

concept of “phronesis” (i.e. “prudence” or “practical wisdom”), this approach to conceiving of rationality is based upon reason as *moral sentiment* and it imagines human agency as being best realized in terms of a practice-based conception of judgment and decision making that aims to produce outcomes based on knowing ‘what to do, when to do it, and how to do it’ (Cameron 2014). In this dissertation, I consider practical wisdom as a form of—or approach to—rationality similar to those mentioned above, but suited to its particular domains and contexts. If I am talking about a *specific conception of rationality* in a given context, I will note this; otherwise I am referring to a suite of approaches of, each with its strengths and weaknesses, each susceptible to the sorts of challenges to cognition that I will outline in chapters to come, and each most concerned with the requirement that individuals be able to make sense of the world in a consistent, valid way that they are capable of sharing with and communicating to others.

Before proceeding, I want to make clear that in this dissertation *I am not concerned with asking whether citizens are “smart enough” for democracy* in the sense of them having the raw intelligence, I.Q., or so on, to think, organize, and decide individually or collectively. Obviously, to some extent, we have managed to do relatively well at all of this despite certain challenges to autonomy and rationality (Heath 2014).<sup>4</sup> I am far less interested in the issue of raw intelligence than in, on the one hand, how we *conceive of* what we are capable of and the expectations that emerge from such a conception, and, on the other hand, the ways in which institutions support, enable, interrupt, or distort our judgments and decisions through exploiting cognitive architectures or tendencies that initially evolved in humans as adaptations

---

<sup>4</sup> The many bizarre, disturbing, and disconcerting events of 2016 notwithstanding.

aimed at other purposes. So, while such structures or tendencies may be related to, or a component of, intelligence as processing power, in this dissertation, I am interested in how individuals *reason* and how they *explain and justify* their reasoning and the conclusions they reach—whatever their “level of intelligence” may be on traditional measures. This approach is defensible primarily because the cognitive distortions and biases, as I will show in chapters three and four, are pervasive across and throughout populations—though, as I note in chapters three and four, there are some distinct intersections that tend to vary with education and socio-economic status and are worth paying special attention to.

## **Epistemic democracy**

In the following pages, I will examine autonomy, rationality, (individual) judgment, and (collective) decision making as they relate to deliberative democracy as a political process aimed at generating judgments that are epistemically valid and authoritative in relation to the norms, facts, preferences, and goals of those who participate in deliberation—what I mean by deliberation as producing epistemically good judgments and decisions. As Warren (2002) argues, deliberation makes public the information and otherwise private reasoning required for reaching collective decisions. This process requires that processes of “challenge, reason-giving, and verification” (192) be undertaken so we can have confidence in the epistemic validity of judgments. As he notes “...the rational validity of a statement—its authority—cannot be separated from the processes that establish this authority in the absence of privileged or objective or independent knowledge” (192). However, as he adds, this requires that a pragmatic consensus is reached (i.e. public understanding) among participants on the

epistemological authority, and, accordingly, the validity, of claims and judgments. As Warren concludes, “validity is a product of *procedures*, suggesting that institutionalized deliberation can establish the epistemic validity of claims and assertions” (193, emphasis mine). This stands in contrast to a correspondence theory of meaning or truth in which validity is a product of “linking” or “matching” the world “as it is” to the world in your head (and by implication in the heads of others).

If deliberative democracy is a theory of democratic decision making concerned with, among other things, producing desirable outcomes based on judgments that are generated by and which reinforce epistemic validity and authority, then rationality and autonomy are essential capacities for those who render judgments. After all, those judgments are required to produce decisions, and both are underwritten by the normative basis of democratic deliberation and the requirement that facts, norms, preferences, motivations, and goals be apprehended, affirmed, publicized, scrutinized, and *publicly known to be* logical and relatively stable. Consequently, to the extent that individual rationality or autonomy is undermined by cognitive limitations, distortions, or biases, epistemic validity and authority are also compromised—and so are the judgments and subsequent decisions generated by deliberative processes. This undermining occurs because the epistemic validity and authority that these judgments and decisions require can only emerge from a pragmatic consensus emerging from exchanges between rational, autonomous agents under (more or less) egalitarian conditions, and within the context of certain personal practices and institutional arrangements that enable good judgment. This reasoning chain is essential to my argument: the undermining of

rationality and autonomy compromise deliberative outputs by short-circuiting validity and authority, and thus producing epistemically suspect decisions (and judgments).

Some judgments in deliberative contexts are reached under conditions in which the rationality and autonomy of participants, and consequently any emergent epistemic validity and authority, are compromised—for instance in cases of misunderstood motivations or distorted interpretations of relevant data/arguments or intentions (a key subject that I explore in chapters three and four). It cannot be said that such judgments reflect the *actual* (considered) will of the participants since it is possible that agents would reach different judgments and outcomes under conditions of complete, or near-complete, autonomy or rationality. As I will argue in chapter two, this concern is particularly acute when inequalities in information processing capacities (e.g. the capacity to critically judge between several alternatives) reflect structural cognitive inequalities (e.g. increased susceptibility to framing effects), both in terms of the extent to which cognitive distortion *affects* particular populations and in which specific populations *are affected*.

### **Nature, culture, and deliberation**

Of particular importance before proceeding is the scope of this dissertation. As I will argue further in chapter three, the root of the fundamental problem addressed herein emerges from a disconnect between, on the one hand, what is required and expected of deliberative citizens (normatively) and institutional arrangements that have emerged from decades of social and political activity (what I am calling, for my own, specific purposes, "culture"), and, on the other hand, the reality of how citizens tend to engage in cognition in deliberative (and other) settings

(what I am calling, again, for my own, specific purposes, "nature") based on our capacities for, among other things, autonomy and rationality. In short: our expectations and our capacities often fail to link up, rendering both our theories and our practice open to critiques about just what we are producing, how, and for whom. To maximize the potential effect of deliberative democracy in generating democratic goods through epistemically good deliberation, *the gap between nature and culture*<sup>5</sup> as it relates to rational, autonomous deliberation must be bridged.

Of course, the cultural milieu in which I am writing this dissertation, and the one it interrogates, is a historical particularity: it comes from a specific time and place, and much of what I address here is far from universal. Maybe none of it is. Even the way we think, in a general sense, is particular. As Henrich et al. (2010) have convincingly demonstrated, styles of cognition in some instances vary significantly from region to region and it is a mistake to state that “we all think” in way X, Y, or Z. Moreover, I do not think it would be such a leap to suggest that ways of thinking vary from historical period to historical period given the variety of concepts, institutions, norms, and imperatives that we see throughout history; however, by and large, enough of the general mechanisms of cognition remain the same despite particular adaptive tendencies in response to upbringing and environment, including our habits and the ways we design the world in which we live that I am able to say something important about how we think and deliberate in contemporary democratic societies. This suggestion is reminiscent of the claim by Xenophanes that while humans tend to anthropomorphize their deities, if horses could draw, they would draw their gods as horses; the expression, in such a

---

<sup>5</sup> My use of “culture” includes institutions in both a formal sense and in the sense of normative behavioural expectations.

case, would be different, but the act would be the same. In the pages that follow, I draw our horses.

Accordingly, this dissertation examines deliberative democracy, autonomy, rationality, epistemology, and cognition within the context of contemporary Western liberal democracies marked by the following characteristics: entrenched, multi-level, and complex governance institutions; the rapid speed and extensive reach of mass communication technologies; persistent disagreement on many issues, both superficial and substantive; and increasingly urban-dwelling populations marked by social, political, cultural, religious, and ethnic diversity. These characteristics are not chosen arbitrarily; rather, they reflect the types of democratic states that deliberative democracy has typically been concerned with, though not exclusively. These phenomena also influence cognition, shaping and directing it through external environmental cues. Of course, none of this *necessarily precludes* the possible portability of much of what I will argue in this dissertation; it does, however, constrain the claims that I am making, situating them in a particular time and place, about particular political systems and those who inhabit them in the early twenty-first century.

Also, as I noted above, not all deliberation is democratic. I am careful to use the term "democratic deliberation" whenever I refer to the *democratic* uses of deliberation—i.e. accessible (to public participants), fair, open (to public scrutiny and review), and regarding questions of a public nature. As He and Warren (2011) find, authoritarian regimes (e.g. China) use quasi-public deliberation for political purposes. And John Rawls (1971) considered the United States Supreme Court to be an exemplary deliberative body. After all, you can have deliberation without democracy, and you can have democracy without deliberation. I am

concerned here with the cases in which there are *both*. I am interested in such cases because I am concerned with the (public) epistemic defense of democratic deliberation as it relates to generating good decisions made by citizens for citizens. In such vastly different settings (e.g. carefully-managed authoritarian political exercises, or in camera court deliberations) there will be various incentives/disincentives, standards, institutions, levels of expertise, resources, and so on would factor into outcomes. Such variance does not mean that my approach to analyzing deliberation cannot be carried out for such deliberations; it could. But that is not my focus in this dissertation.

### **Affect and autonomy in political philosophy**

To fully understand how and why citizens deliberate (or fail to deliberate) well (i.e. rationally and autonomously), we must consider human psychology. Traditionally, however, political theorists in general and democratic theorists in particular have paid only limited attention to the impact of a-rational cognition on rational, autonomous political judgments and decisions, despite the fact that these phenomena can contribute significantly to shaping each. More specifically, to the extent that a-rational cognitive processes condition, alter, or direct autonomous judgment, they may significantly impact real-world political outcomes, both directly in the moment of judgment or decision and in the future by forging a familiar path that subsequent judgments and decisions may traverse (i.e. via path dependency). Understanding how such processes and stimuli affect autonomous, rational judgment in deliberative settings will help reveal a full portrait of the citizen as a deliberative agent. It may also yield some

effective personal and institutional tactics and strategies for improving judgment in the future—which are the focus of the second half of this dissertation.

While theorists of politics have yet to provide a full account of how a-rational cognitive processes affect rational, autonomous judgment in deliberative settings, several political theorists and philosophers have provided accounts of the impact of affect<sup>6</sup> on political judgments and decisions. In fact, stretching back at least as far as Plato, political theory and philosophy have folded affect into the question of what, if anything, makes a human capable of rational, autonomous judgment and under which conditions the exercises of each capacity is possible. And while such enquiries were not always expressly political, for these thinkers, understanding the origins of the causes of one’s judgment has been essential to account for agency and freedom, since such origins may have implications for how politics should or should not be undertaken and under which conditions. Ultimately, in this dissertation I aim to build upon these past discussions of rationality and autonomy, updating previous accounts with new findings from social and political psychology, and placing the matter within the context of deliberative democracy as a theory of how we can generate epistemically good judgments and decisions.

But first, we return to the beginning. As noted, Plato was one of the first philosophers—perhaps *the* first—to examine how affect affects judgment. In *The Republic*, Plato divides and orders the human soul (or mind) into three elemental components, placing the logical or reasoning element at the top of the hierarchy, and the appetitive and spirited (including

---

<sup>6</sup> In this dissertation I assume that affect includes emotion and feeling. Traditionally, political theorists have often used emotion/feeling/affect interchangeably, and in these pages I will do the same. In the few cases where I need to distinguish between the terms, I do so expressly.

emotional) elements below the rational element, since they are in need of regulation and control in order for a human to be well-ordered and capable of rational, autonomous, enlightened judgment (2004 [c.380 BCE]). Aristotle, a student of Plato's as well as his critical interlocutor, was also wary of the potentially deleterious impact of affect run amuck, arguing in *The Nicomachean Ethics* that the incontinent man—one who is without self-control, such as one ruled by a-rational forces—"acts from desire but not from choice" and therefore does not act entirely autonomously (2004 [c.350 BCE]: 1111b14-16). He even goes so far as to determine that the best (i.e. happiest) activity of human life is the rational contemplation of the universe and its eternal truths (2004: Book X). Nonetheless, Aristotle makes space for affect in human life, arguing that it is an essential element of well-being, and even useful in judgment insofar as it may guide judgment, but only if its role is understood, moderated, and its use is kept in check (within the mean between extremes).

Plato and Aristotle initiated the tradition in political theory and philosophy of asserting reason as the prime and most effective capacity that makes a human being capable, in a cognitive sense, of undertaking political activity autonomously, thus linking the two capacities. Whatever the place or necessity of affect to the process of thought, it was seen, at best, as suspect—and likely to be a distorting phenomenon. The subsequent history of the study of politics in the West stayed the course set by Plato and Aristotle, remaining skeptical of affect and its effects on thought and behaviour, if not quite universally denigrating it. One of the most pronounced and sustained critiques against affect and its place in cognition is offered by Spinoza, who, in *The Ethics* (2000 [1677]), argues that to the extent that human beings act from emotion (a term Spinoza uses explicitly and carefully) they act passively and their autonomy is

significantly diminished. Agents who are influenced by emotions, especially those that are not fully understood, are acting passively. Passive action occurs when an agent is unaware of the origins of her actions, and since freedom for Spinoza is active judgment, then to the extent that one is unaware of the origins of her actions, she is unfree.

Roger Scruton (2002) summarizes Spinoza's argument nicely, noting that for him "freedom is not freedom from necessity, but the consciousness of necessity" (64). Thus, Spinoza locates an agent's autonomy in her capacity to know the origins of her judgment. In so doing he anticipates—perhaps initiates—what would become the apotheosis of reason and rationality in modernity: the rise of Kantian epistemology. In this dissertation, the conceptions of autonomy offered by Spinoza and Kant will play an important role in how I define autonomy and conceive of its function vis-à-vis cognition and political judgment; these conceptions of autonomy underwrite my argument that good political judgments require extensive and accurate knowledge of their origins *in order to guarantee that a judgment is good*, so that it can be communicated to—and in some way held in common with—others, even if that is only for the purposes of disagreeing with one another.

Recall that Kant (2012 [1785]), for all his revolutionary reordering of our understanding of the nature and expression of human epistemological capacity, argues a point similar to Spinoza's—a point that is at the centre of my core argument in this dissertation. Kant makes the distinction between *autarchy*, which is the ability to make choices for oneself, whatever the motivation one may have (known or unknown), and *autonomy*, which requires that the agent can rationally reflect upon and give reasons for his judgment. In conceiving of this distinction, Kant implicitly set a standard for judgment that entrenched rationality as the ultimate

guarantor of sound, autonomous judgment (including autonomy as necessary for *communicating reasons for that judgment*). Judgment for Kant requires that the agent engage his capacity for practical reasoning (cf. Aristotle): only rational reflection within the constraints of the macro-structures of human cognition and voluntary subjection to rational laws can yield autonomous judgment and action; if such reflection and subjection are absent, an individual is considered to be acting heteronomously (e.g. acting from desire or fear, or some other a-rational determining force). Again, as in Spinoza's conceptions of autonomy and judgment, to the extent that an individual is unaware of the reasons for their judgment, or unable control their emotion when making a judgment, their autonomy is diminished, if not eliminated altogether.

Returning to deliberative democracy, the tradition of theories of democratic deliberation have remained thoroughly, if not universally, Kantian. Reasons given by an agent are expected to reflect rational considerations (even if they contain some emotional element or are presented emotionally) and are supposed to be anchored in goals and preferences about which the agent can reason and communicate in the hopes of generated shared understanding (i.e. validity) through communication with others. Thus, within the cognitive processes of an agent, a rational, autonomous link is presupposed: stretching and connecting from rational, autonomous (personal) reasoning to shared (collective) understanding. The universal presence and proper functioning of this link in agents are, however, doubtful. In fact, it may not even be the case that this relationship is *commonly* present when agents deliberate, as I will argue throughout this dissertation.

Perhaps the first great doubter in this regard was Hume. This is because Hume de-emphasizes reason as a capacity with significant causal force and denies that it can be the primary source of either morality or judgment; instead, he introduces affect (“passions”) into the equation. He defines the will as “...nothing but *the internal impression we feel and are conscious of, when we knowingly give rise to any new motion of our body, or a new perception of our mind*” (Hume 1975 [1738]: 399). He adds that “Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them” (Hume 1975 [1738]: 415). As Lara Denis (2012) notes regarding the motivational roots of our actions and judgments (which is worth quoting at length), according to Hume

[a]bstract (or demonstrative) reasoning, which involves a priori inferences and judgments about relations of ideas, cannot influence the will, but only assist us in our pursuit of an end we already have (e.g., if mathematical calculations would facilitate our achievement of our end). Probable (or causal) reasoning helps us discover cause and effect relations among objects of experience conducive to the realization of pre-selected ends, but such information about cause and effect can never motivate action on its own...In order to be motivated to act, we must first anticipate pleasure or pain from something. That anticipated pleasure or pain gives rise to feelings of desire or aversion for the object in question.

Probable reasoning allows us to discern the causes of this object; our positive or negative feelings about the object then spread to the causes of it; and we are then motivated to pursue or to avoid them... (no page reference)

This conception of the process of reasoning includes the act judgment and, accordingly, if it is accurate as an account of human cognition, renders elusive the kind of autonomous judgment required for deliberative democracy to function as a theory of how epistemically good decisions might be made. Why? Because if Hume is right, the requirement that public reasons be given in concert with others in an attempt to generate preferences and reach decisions in a deliberation is potentially reduced to the mere aggregation of affect-driven desires coated with the gloss of ex post facto (non-conscious) rationalization.

This outcome is, of course, the epistemically worst-case scenario for proponents of deliberation. The concerns raised about our theoretical goals as deliberative democrats as they relate to epistemic validity, emerging from autonomous judgment, are thoroughly Kantian and are based upon a high standard of judgment and decision making. Nonetheless, doubts about the extent to which such validity is possible, emerging from work in social and political psychology, which are closer to the Humean tradition, must be taken seriously—though they should not be overstated—if we are to produce the best arguments we can for the importance of deliberative democracy. To put this dissertation in its historical place, then, it may be fair to say, that it is a contemporary attempt to continue the debate between these grand ideas about autonomy, rationality, judgment, and decision-making in the history of philosophy, with an eye to both grounding and extending contemporary debates on these subjects.

## **Cognitive limitations, distortions, and bias in political contexts**

Notwithstanding the divergent theoretical positions on the need for and nature of autonomy and rationality, observational and experimental work in social and political psychology has revealed the extent to which human judgment and the capacities of autonomy and rationality that underwrite them can be compromised in political (and so many other) contexts. Kuklinski and Quirk (2000) summarize the problem well: cognitive biases and distortions have been hard-wired into the human brain (e.g. the tendency to rapidly categorize stimuli as friendly or threatening without deeper interrogation, or stereotyping) through a long history of evolution; and while, as the authors suggest, such distortions are likely to have served a function at some point in our history (and perhaps still do in some ways), they present serious challenges to contemporary democratic citizenship and the normative expectations this citizenship includes. As they conclude: “researchers should not presume that any feature of human cognition is well adapted to the tasks of citizen” (Kuklinski and Quirk 2000: 165). In sum: we have evolved for life in a prior evolutionary age and, to once again quote Kuklinski and Quirk (2000: 166) in their eloquent skepticism: “...we cannot assume that the cognitive processes people use in making a particular political judgment are well adapted to that use. To the contrary, if close observers see such processes as irrational or misleading, they probably are.” This conclusion points to the argument I made earlier—and will make more extensively in the next two chapters—that there is a divergence between ordinary individual capacities *as they currently tend to be exercised* (nature) and what is nonetheless expected from those individuals.

There is a myriad of experimental and observational data that support this scepticism. As early as the 1950s, systematic research emerged in political psychology that demonstrate

that political judgments and decisions—such as vote choice or opinion/preference formation and expression—were at least partially determined by factors that were not themselves expressly political or even rational. Berelson et al. (1954) provide one of the earliest and best iterations of the choice model, arguing that political socialization—learned primarily from the family—helps to explain the continuity of voting patterns and attitudes throughout time, casting vote preference as a learned behaviour that is transferred between social groups and then repeated over time, like a habit. By the 1960s, Converse (1964) had emerged to present an even more dire evaluation of (non-elite) voters, suggesting that they tend not to have coherent belief systems and often fail to evaluate candidates or parties on the left/right political spectrum, though this claim is disputed (see Jost et al. 2006). Voters, Converse concludes, are not typically ideological, and many tend to change their mind seemingly at random on certain issues. Reminiscent of this position, though somewhat more sanguine about voter competence, Zaller (1992) argues that citizens draw on competing and sometimes conflicting considerations when making political decisions or forming opinions; the decision that is ultimately taken by an individual comprises a sampling of the most pertinent (and often the most recent) information found “at the top of their head” at the time when they are required to make a decision or pass judgment. In Zaller’s research, what emerges is an understanding of the voter as having *something like a political attitude*, though it remains variable and unstable.

Regarding political information processing, researchers have discovered several phenomena that call into question the extent to which individuals are capable of rational, autonomous judgment. For instance, framing and priming effects have been found to occur in subjects who process political information, causing them to change the weight they place on

the significance of information based on the frequency and order of mentions of the subject (Chong and Druckman 2007; Iyengar et al. 1982). However, Lenz (2009), argues that the phenomenon is in fact evidence of learning and not manipulation *in real-world observational findings*.<sup>7</sup> Regarding elections, Ballew and Todorov (2007) find that rapid judgments of the perceived competence of candidates in gubernatorial elections predicted outcomes of the race based on 100ms of exposure to the faces of the candidates (who were previously unknown to the subjects); while researchers have not isolated the mechanism at work in these predictions, Ballew and Todorov suggest that the facial expressions of the candidates—demonstrating (implied) competence or not—may play a role in determining who the subject predicts as the winner. Each of these instances is a challenge to contemporary accounts of advanced democratic citizenship, given that they undermine the normative expectation of sustained, rational reflection and autonomy in judgment and decision making, and they suggest common cases in which a more rapid, intuitive judgment is at work (which is prone to error, though not universally so).<sup>8</sup>

Psychologists have pressed matters further on this front, examining how unsettling stimuli, including being read a surrealist story, being subtly reminded of death, or playing cards

---

<sup>7</sup> Lab findings from priming experiments, however, cannot be explained as a product of learning.

<sup>8</sup> Priming effects have been shown to, generally, have a temporally limited and transient effect. However, this does not mean that priming effects are insignificant. For one, repeated priming that leads subjects in a particular direction may have long term effects. Also, decisions tend to be made by individuals under conditions in which they are primed in or around the moment—e.g. at the polls, during discussions, while consuming news and forming opinion. The accumulation of decision outcomes based on priming in some particular direction—e.g. framing a war as a battle for civilization or making a news story seem more salient by continued coverage—may produce broader outcomes of significance. Thus priming may be significance via repetition (at the site of the individual) or accumulation (across individuals).

with the suit colours swapped, affect subjects' behaviour. In laboratory studies, subjects exposed to unsettling stimuli tended to non-consciously employ the coping mechanism of affirming alternative belief systems (and overcompensating by behaving in more extreme ways when they did). Subjects presented with discordant stimuli tended to be more likely to assert a "law and order" belief system, becoming more likely to assign more significant punishments to a hypothetical offender than subjects who were not presented with such stimuli (Heine et al. 2006; Proulx and Heine 2006; Proulx et al. 2010). The force of these effects seems to rely on non-conscious motivation via implicit messaging—whether or not those messages are expressly intended or not. Mendelberg (2001) has shown the significant effect of implicit messaging surrounding race politics in the United States, finding that race-baiting is generally effective when it circumvents established race-norms of equality among white voters by subconsciously activating feelings of resentment and negative predispositions towards African Americans by pairing certain stimuli (e.g. the image of a "frightening" black man and foreboding music).

Not only do implicit messages impact judgments and decisions, undermining the exercise of autonomy and rationality in the process, attempts at defending such judgments may be compromised by further cognitive limitations and distortions, thus even further casting doubt on the possibility of autonomous judgment. As Sniderman et al. (1986) demonstrate in their study of Americans' support for welfare for African Americans, the reasoning chains of some voters may not proceed rationally and serially; rather, conclusions justified *ex post facto* with (rational) reasons may, in fact, have their origins in affective presuppositions, prejudicial or otherwise (e.g. masquerading racism as concerns about moral hazard). Such rationalizations

deviate from the standard of autonomous judgment expected and required by, among others, proponents of deliberative democracy.

Many of these findings suggest that we cannot always assume the existence of a common, sustained, and refined capacity for citizens to express rational, autonomous judgment when it comes to political matters—especially when that requires agents to know the sources of their judgment (or behaviour). Whether those sources of judgment are conscious or not, political or not, or rational or not, they must be accounted for if judgment is to be considered autonomous and if an individual is to arrive at a generally rational judgment based on reasons they can share with others to generate validity in the context of democratic deliberation. While past standards for understanding the place of a-rational cognition and the impact of certain types of stimuli on judgment were crude, today we have new data, methodologies, and a much more sophisticated and empirically supported sense of many (though probably not all) of the factors that impact an individual's judgment and a group's decision. Now, if we wish to come to a theory of deliberative democracy, defended on epistemic grounds—grounded in the capacity for autonomous and rational exchanges of reasons and arguments and the possibility of shared understanding—then we must take up the ancient question of what are the sources of autonomy and rationality when it comes to judgment and decision making; moreover, we must adapt this question and its implications in light of new findings about the nature of human cognition as it relates to democratic deliberation—which is, in part, what I set out to do in this dissertation.

Some scholars of political psychology have already begun to do this. In their exhaustive study of how voters think, Lodge and Taber (2013) argue that the expectation that voters (and

by implication, citizens more generally) typically turn to intentional rational evaluation to generate deliberative political judgments is grossly overstated. As they claim,

[f]ar more common [than intentional rational evaluation], we believe, will be the reverse causal pathway from evaluation to deliberation. This *rationalization* hypothesis...asserts that the causal pathways...that travel through unconscious affect, and in particular the affect-driven evaluation processes, cause most of our deliberation about politics.... Evidence is accumulating...that attitudes and behavioral intentions—even behavior itself—arise from automatic, uncontrolled processes and are often set before we begin seriously “thinking” about them. This the case, deliberation serves to rationalize rather than cause (21, emphasis in original).

What we are left with, then, is what Lodge and Taber refer to as rationalizing, rather than rational citizens. These citizens are guided by affect-driven, dual-process modes (I address dual-process modes in depth in chapter three) of thinking and reasoning. They tend to rely on the second stream of cognition in the dual-process model: the rapid, automatic, non-conscious,<sup>9</sup>

---

<sup>9</sup> Throughout this dissertation I will use the word “non-conscious” rather than “unconscious” wherever possible—that is to say, whenever I am not directly quoting an author who has chosen the latter. This is for the simple reason that I wish to avoid any association with common conceptions of the Freudian unconscious (hidden but vague “drives,” sexualized motives, the tripartite division of personality), which, while broadly relevant to the subject, is merely adjacent to the project at hand. While I will discuss drives that are outside of awareness, the data and theories I draw on are very different, and more recent and sophisticated than those available to or used by Freud.

affective stream, rather than the slower, deliberate, and rational first stream (see also Kahneman 2011). The authors summarize nicely the emergent doubts about citizens' capacity for rational, autonomous judgment in their conclusion that "affect precedes and contextualizes cognition" (27). I will revisit this conception of political cognition in chapter four.

## **Models from social and political psychology**

Past and recent accounts as well as models of cognition offered by students of social and political psychology tend to support the claims made by scholars such as Lodge, Taber, Kuklinski, Quirk, Sears, and Marcus, that non-conscious, affect-driven thinking often drives citizen judgment. In this dissertation, I will draw primarily on four models of cognition developed by top researchers in the field of social and political psychology and I will contextualize with reference to system 1 and system 2 cognition (Kahneman 2011). The system 1/system 2 division is, more broadly, a dual-process model of cognition in which system 1 represents fast, automatic, emotional, non-conscious cognition, while system 2 is characterized by slow, rational, conscious thought. In chapter three I will explain which elements or modes of each model I use falls into system 1 and which into system 2, and under what circumstances; and I will link each back to rationality, autonomy, and the generation of epistemically good judgments and decisions within the context of democratic deliberation by highlighting the challenges (and potential opportunities) presented by each.

While several scholars have contributed to the development of each of these models, I will refer to them primarily with reference to their originators and best-known proponents, though I will also engage with their critics where it is useful to do so. The models are Richard E.

Petty and John T. Cacioppo's *elaboration likelihood model*, John T. Jost's model of *system justification*, Jonathan Haidt's *social intuitionism*, and John Bargh and Tanya Chartrand's *automaticity*. While each of these models, broadly stated, offers an independent account of the nature and processes of particular modes of cognition as they relate to human behaviour and judgment, they are not mutually exclusive. Indeed, part of the appeal of these models is that they are consistent with one another and also compatible with the organizing systems/dual-process model mentioned above. I will also discuss motivated reasoning on its own, in chapter four—which is the kind of model used by Lodge and Taber, noted above. Motivated reasoning has been given its own chapter because it stands apart from the other models. Why? Because motivated reasoning draws more on explicit *motivational* considerations than the other models, which rely more heavily on *cognitive* considerations. Still, there is overlap between motivated reasoning and the other models, which I will note and discuss when appropriate.

I have chosen to explore these four theories for five reasons. First, they enjoy theoretical compatibility with one another; second, they enjoy the status of leading theories in the fields of social and political psychology; third, they are plausible, cogent, and rely on rigorously collected and theorized data; four, they are intuitively appealing (the irony is not lost on me here); and fifth, they correspond with emerging theories of behaviour and judgment within political psychology and the data supporting those theories. Accordingly, I have chosen these to include these models together as elements of a broader approach to understanding judgment and behaviour through the lens of non-conscious political cognition. On balance, these are plausible and representative models—and the particular findings I will use to explore and support them are just as plausible and representative. Moreover, these models are capable

of coherently absorbing and accounting for data from not only social psychology but also political psychology and cognitive neuroscience. The macro-level coherence, representativeness, and cogency of these models is what, ultimately, renders them useful and very well fitted for this study of rationality, autonomy, and cognition as they related to epistemically good (or not) democratic deliberation in liberal democracies.

I will examine these models in chapters three and four, so I will only briefly sketch each of them out here. First, Petty and Cacioppo's elaboration likelihood model suggests that attitude change occurs through one of two cognitive processing routes—central (i.e. system 2) or peripheral (i.e. system 1). Second, Jost's model of system justification asserts that certain psychological motives and processes—for instance, the desire for stability or the fear of death—are causally linked to ideological predispositions (see Jost 2006; Jost and Amodio 2012; Jost et al. 2003). Third, Haidt's social intuitionist model suggests that personal moral reasoning does not cause moral judgment. Instead, such judgments emerge from interpersonal processes and are perceived, rather than rationally discovered. The moral reasoning that gives ideational foundation to these judgments appears only after the fact (Haidt 2001). Finally, Bargh and Chartrand's model of automaticity suggests that our capacity for intentional and rational judgment is limited by automatic perceptual direction taken from the external environment, which directs behaviour and judgment (Bargh and Chartrand 1999). Moreover, action precedes reflection; both external stimuli and internal mental processes contributing to behaviour and judgment precede rational reflection upon—or even apprehension of—behavioural outcomes and judgments; tendencies towards certain sorts of actions in these areas develop over time, creating general, automatic, and non-conscious patterns of behaviour and judgment in citizens.

Conscious apprehension of such activity emerges later, acting as a “gatekeeper and sense maker” (Bargh and Morsella 2008)—and what, as I will argue, can be characterized as rationalizations or pre-deliberative motivations that become expressed as preferences, reasons, and judgments when individuals deliberate.

## **Chapter outline**

As I have noted, the primary question I address in this dissertation is: How do a-rational cognitive processes affect of the possibility of rational, autonomous judgment and the production of epistemic validity in deliberative settings? The secondary question that I address is: How can political theory and political science respond to the effects of such processes and stimuli in a way that underwrites capacities of rationality and autonomy, and the production of good judgments and decisions from democratic deliberation? These questions emerge in response to an attempt to bridge the gap between what is normatively expected of those who deliberate in a liberal democracy, and what they tend to be able to deliver.

To address these questions in a way that relates them to one another, this dissertation is separated into two parts. The first part (chapters two, three, and four) makes the argument that democratic deliberation requires a relatively high-level capacity rationality and autonomy, and that certain a-rational cognitive processes currently pose a serious and persistent threat to this requirement by undermining the human capacity to give reasons for judgments that reflect real motivations, eroding the capacity for deliberation to generate epistemically good judgments and decisions.

To this end, in chapter two I argue that autonomy is the keystone capacity required for deliberative democracy to function as an approach to generating epistemically valid and authoritative judgments and decisions among citizens. Recall that the focus on autonomy over rationality is that when the former is present to a high degree, the products of the latter can be better interrogated and understood. Deliberative democracy is founded on reason-giving and the force of argument as a way of reaching legitimate and just political outcomes; for these to be achieved, the sources of judgments must be known and agents must have confidence that the facts, norms, motivations, and desires that they adhere to or express are not the products of cognitive distortion or bias. Otherwise, the moral grounding of the theory of deliberation as one based on good faith attempts to approach decisions based on the force of reasons and the better argument is eroded by the influential presence of possibly irrelevant determinants and important factors that might remain outside the realm of interrogation and the democratic goods produced by epistemically valid reasons may be undermined or lost. Moreover, I argue, building on Kant's distinction between *autarchy* and *autonomy*, that to satisfy the requirements for a judgment to be autonomous, not only must an individual be able to give reasons for that judgment, those reasons must closely reflect their *actual motivations* for holding them. That requirement is meant to ensure that reasons are not generated heteronomously—or, to be blunt: their reasons must be accurate and their own. This requirement aligns closely with Christman's conception of autonomy, the definition of that concept that I will use in this dissertation and will explore in depth in chapter two. According to Christman, autonomy is a capacity in which "...the influences and conditions that give rise to the desire [or preference or intention] were factors that the agent approved of or did not resist, or would not have resisted

had she attended to them, and that this judgment was or would have been made in a *minimally rational, non-self-deceived manner*" (22, emphasis mine).

In chapter three I argue that the capacity of autonomy required for deliberative democracy, defended in chapter two, is undermined in some cases by certain a-rational cognitive processes, and that these processes also undermine capacities for rational judgment. I frame the discussion as being rooted in a disconnect between what we expect and require in our culture and what we tend to be able to do to given our evolutionary cognitive limitations. By culture I mean the social, political, and technological institutions that we have developed (including their normative content, such as ideas about how we ought to organize our institutions and how we ought to behave towards one another); by nature, I mean the central tendencies and limits of our biological and psychological capacities as they have developed over the course of our time as a species. I then present findings from social and political psychology and cognitive science that demonstrate the ways in which rational, autonomous judgment is susceptible to influence—or even determination—by apolitical forces or phenomena that have traditionally been considered irrelevant to arriving at political judgments in deliberative settings. Specifically, as noted above, I use four generally accepted models of human cognition organized under system 1/system 2 distinctions: Petty and Cacioppo's elaboration likelihood model, Jost's system justification, Haidt's social intuitionism, and Bargh and Chartrand's automaticity. I explain why I have chosen these models and how the threats to autonomous judgment they raise undermine normative arguments in favour of deliberative democracy as both a moral theory founded on rational, autonomous reason giving and the force of the better

argument, and as an approach to generating epistemically good or correct judgments and decisions.

I also argue that given the strength of these threats to deliberative democracy as a political process for reaching legitimate and just decisions based on such epistemically valid and authoritative facts, norms, and preferences, and given our current conception of rationality and autonomy as they relate to judgment in deliberative settings, theorists of deliberative democracy who wish to address this challenge are presented with four general approaches in response: ignore, revise, reject, or respond. Scholars of deliberation can ignore the problem altogether and accept the potentially flawed nature of the judgments and decisions that follow. They can revise deliberative democracy as a moral theory, altering its moral foundation (i.e. honest reason giving and the force of the better argument) or the requirements for what counts as a good reason or a good argument. They can reject deliberative democracy entirely and adopt a more explicitly, traditionally “political” (i.e. realist, aggregative) approach to arriving at political decisions; this substitute approach would need to be one in which concerns about the particular motivations and determinants of judgment were made secondary to (or perhaps eliminated entirely from) outcomes. Or they can, as I argue they should, seek to develop approaches towards designing institutional set-ups and personal practices to be deployed within a deliberative system aimed at leveraging our cognitive capacities in such a way that we can reach rational, autonomous judgments through deliberation.

This approach requires an account of what counts as rational, autonomous judgment in such settings so that we may know if the requisite standards are met; it also requires a statement of the conceptual and practical apparatuses needed to meet such a standard—each

of which I cover in part two of this dissertation. While this may seem like a revision to the theory of what counts as a good decision or judgment, it is not; the requirement of autonomy offered in accounts of deliberative democracy is maintained, as are the guiding ideal requirements that reasons are given and that the force of the better argument carries the day. Instead, this approach adds a new *evaluative* criterion (i.e. that those who arrive at judgments are aware of all the relevant factors that determine those judgments) and suggests ways that this standard might be met by those who deliberate.

In chapter four I examine the cognitive phenomenon of motivated reasoning—primarily, though not exclusively, through Lodge and Taber’s model—as a specific motivational threat to political judgments and autonomous deliberation, linking that model back to the four general models explored in chapter three and examining how it relates to deliberative democracy. Motivated reasoning is a phenomenon consistent with each of the four models, especially when such reasoning occurs under the system 1 mode of cognition (though it is also compatible with system 2). Additionally, in this chapter I define and examine the related concepts of hot cognition and reasoning chains and link them to democratic deliberation. I explore critiques of the models I discuss.

In the second part of this dissertation (chapters five to seven), I present a general approach to rethinking autonomy in deliberative settings as well as specific responses to the challenge of a-rational cognition to rationality and autonomy. I suggest that to address the challenges I raise, we need to both add new tools to the deliberative democracy toolkit and mobilize and adapt existing tools. I also discuss the specific steps required to move from

democratic deliberation as it currently is practiced to a sort of democratic deliberation that can specifically respond to the challenges outlined in this dissertation.

In chapter five I ask whether we can design deliberative practices and institutional arrangements in such a way that underwrites and enhances our capacities for rationality autonomy. I introduce and develop the concept of institutional deliberative autonomy (IDA), into which I fold concerns for both capacities of rationality and autonomy. IDA is a concept developed for thinking through whether rational, autonomous judgments are reached in deliberative settings. Accordingly, IDA includes an account of what counts as rational, autonomous judgment in deliberative settings—an account that is consistent with a robust theory of autonomy. This concept is not merely an evaluative concept, but also a regulative ideal meant to guide deliberative democratic institutional designs and evaluations of participants and outcomes. IDA does not track the status of a binary state (rational/autonomous or not), but rather within it I imagine a scale for autonomy that ranges from maximal autonomy (i.e. the absence of any determining factors save for rational judgment—a theoretical high-point that is neither achievable, nor necessarily desirable) to total automaticity (i.e. the full determination of one’s judgment by factors entirely outside of one’s awareness).

In this chapter I also introduce and interrogate four concepts for thinking about autonomous, rational deliberation might be brought about. The first concept is *iteration*: repeated discussions or exercises, similarly structured over the course of a period, and designed to develop a stable central tendency and to wash out individual instances of random cognitive

distortions. This approach, however, does not apply to structural distortions; for that, other concepts are required.

To that end, the second concept I introduce is *a-rational receptivity*, which refers to personal practices aimed at attempting to, as best as possible, bring to the surface hidden, partial, or manifest moods, feelings, emotions, motivations, fears, anxieties, and so forth during deliberation. More specifically, a-rational receptivity refers to a group-wide disposition aimed at encouraging the exploration of an individual's a-rational state. Such an element is also an institutional feature because it is built into the structure of the deliberation and employed as necessary; however, its operating mechanisms are individual practices of self-interrogation combined with group-wide receptivity to such interrogations. Returning to Kant: autonomy requires that individuals have both judgments and reasons for them; I hasten to add that autonomy requires that individuals have reasons for judgments, which would include the emotions, feelings, and moods that play a role in generating and underwriting those judgments.<sup>10</sup>

The third concept I look at is *cognitive diversity*.<sup>11</sup> This concept reflects an attempt to prevent individuals from coasting through deliberations by relying merely or mostly on automatic, low-level or non-conscious heuristics or gut-level guidance. While these automatic

---

<sup>10</sup> Though it is presently outside of the scope of this dissertation, a-rational receptivity may be a concept useful for bridging the gap between theorists of deliberation who rely on traditional understandings of autonomy and reasoning, and theorists of difference who are critical of the potential exclusivity and limits of this approach (see Benhabib 1992, 1996; Bohman and Rehg 1997; Young 1990).

<sup>11</sup> Cognitive diversity does not necessarily require a diversity of identities, only a diversity of approaches towards cognition and its related practices. That said, to the extent that differences in identity correlate with cognitive diversity, the two can be thought of as bound up together, though this is not logically necessary.

processes may often be useful for navigating a complex and variable world, the kind of attention and reflection that deliberation requires is usually poorly served by this approach. Cognitive diversity may hold part of the answer to getting subjects out of this state. As defined by Landemore (2013), building from Hong and Page's (2004) conception, cognitive diversity "refers to a diversity of ways of seeing the world, interpreting problems in it, and working out solutions to these problems. The concept denotes more specifically a diversity of perspectives...interpretations...heuristics...and predictive models" (see also Page 2007). In deliberative contexts, the presence of a variety of individuals with different cognitive styles might not only improve decision making (Page 2007) and the quality of deliberation (Landemore 2013), but that presence of diverse ways of thinking may also offer a cognitive jolt to individuals who might otherwise rely heavily on the low-resource flow of automaticity when processing information and reaching judgments.

My fourth and final concept is *targeted motivation*. This practice indicates an attempt to engage participants in central-processing (i.e. deliberate, rational reflection rather than low-level, automatic processing) by highlighting the relevance and importance of the issue or issues at hand. Targeted motivation may assist in shifting subjects' attention towards the subject matter, thus maximizing the likelihood that they will scrutinize the data and arguments presented to them. Such scrutiny, in theory, would open individuals up to iteration and a-rational receptivity, as well as traditional approaches reasoning, by raising the cognitive stakes of the deliberation. Ultimately, targeted motivation alone may not work on all deeply embedded processes linked to cognitive distortion; it may, however, minimize the impact of

some of the superficial effects of automaticity and may enable deeper tactics, such as a-rational receptivity, to work.

In chapter six I move away slightly from the individual in the context of democratic deliberation and look at the *broader social and political* institutions that frame democratic deliberation. I compare different understandings of what an institution is and examine how both such conceptions of institutions and the institutions themselves interact with those who deliberate. I also ask how these institutions might be understood and reformed in a way that enhances democratic deliberation. I look specifically at the media and attitude formation, political parties and partisanship, electoral systems and elections, and ideology and hegemony. My primary focus in that chapter is how institutions often come to encourage poor thinking, judgment, and decision making. This focus sets up chapter seven—which is an examination of deliberative systems.

In chapter seven I discuss how an understanding of deliberative systems might help us to respond to the challenges of both individually limited and distorted cognition and institutional arrangements that undermine our best efforts at producing epistemically good democratic deliberation. I argue that good democratic deliberative systems design—along with the aforementioned institutional reform and changes to individual practices—can help generate better deliberation despite challenges from the sorts of cognitive distortion explained by the four models I examine in chapter three and motivated reasoning. This chapter pulls together several threads: from the individual and social and political institutions, up to the level of a democratic deliberative system that goes beyond mere one-off deliberative events and

entrenches deliberative activities as complementary approaches to democratic judgment and decision making in a political system.

In chapter eight I conclude the dissertation by recapitulating how each of the arguments raised in the dissertation connects to one another and examining the problem of cognitive distortion as it relates to undermining the rationality and autonomy required for reaching epistemic validity in deliberative contexts. I also revisit the conceptual approaches I have offered in response to these challenges, further arguing that these approaches require testing in real-world deliberative contexts to determine whether they will be able to address the challenges to rationality and autonomy discussed throughout the dissertation. I end with a discussion of how the cognitive sciences might be further integrated with political science and political theory in the future.

## Chapter 2: The Epistemic argument for democratic deliberation and the need for autonomy

Liberal democracy refers to a political system predicated on, among other foundational elements, the related premises that individuals are the best judges of their own interests and their own preferences (Dahl 1989; Raz 1986). Consequently, it follows that those within such an association ought to be able to choose freely, at least in a general sense, how that association operates (Habermas 1996; Held 1995). Given these premises, it is important to carefully determine by what means members of such a body should make decisions that will have an impact on their lives and the lives of others. If each citizen were an island unto herself, then it would scarcely matter to anyone else what that individual wanted or why she wanted it; nor would it matter what she did. However, liberal democracies are populated by individuals whose interests and preferences diverge from—and often conflict with—one another and within which the interests and preferences often affect the lives of others. These divergences must be addressed if legitimacy, order, good government, and just policies and laws are to be established and maintained. Accordingly, it is necessary to ask *precisely how* this is to be done and to develop a procedure (or procedures)<sup>12</sup> For decision making that is both fair (i.e. acceptable to reasonable participants and impartial towards all who deliberate) and effective at

---

12 To say a bit more, for this dissertation I conceive of fairness in limited terms that are specifically related to democratic deliberation. For my purposes, deliberation is fair if the terms of the deliberation are i) generally acceptable to all those included or affected; ii) those terms are followed during deliberations; and iii) the deliberation—notwithstanding whatever other terms have been set—is impartial among those who participate in such a way that no single individual or their perspective is *a priori* privileged over and above that of another. Of course, during the deliberation an individual or the collected body may *choose to* favour a perspective, but this should not be taken as granted before deliberation has begun nor should it be chosen for individuals or the collected body externally.

producing outcomes that link individual and collective interests and preferences with outcomes in the formal political sphere.

Deliberative democrats advocate a participatory system in which citizens (or their representatives) exchange reasons with one another to generate political judgments or decisions<sup>13</sup> that may be used to generate recommendations, laws, or policies that will be enacted by decision makers or put to the broader population for approval. Rather than suggesting that decisions are best made by merely aggregating pre-existing preferences—a blunt, if standard, majoritarian approach—deliberative democrats argue that the deliberative approach to democracy yields more legitimate, just, and valid outcomes (Warren 2002). Theories of deliberative democracy assume that those who participate in judgment and decision making are autonomous (i.e. they are the originator of their actions and preferences, and can give an account of why they act in such a way or hold such preferences) and rational (i.e. they can more-or-less accurately discern facts about the world in a consistent way, that they can communicate them to others effectively, and that they are open to being persuaded by reasons). However, these standards for deliberation, *especially the standard of autonomy*, leave a gap between expectations about one's ability to establish natural and normative facts about the world, and the extent to which, in practice, individuals tend to be fully aware of *why* and *how* they arrive at a judgment (and how those collective judgments generate decisions).

My interest in this dissertation is in the assumption underlying the requirements of autonomy in reason giving, which should also be taken to imply an extended sense of self-

---

13 Recall that for the purposes of this dissertation, "judgment" refers to individual conclusions arrived at after a period of (perhaps very brief) consideration, while "decision" refers to a binding or semi-binding conclusion reached by a particular group (e.g. a deliberative body).

ownership and self-direction: namely, that one is *mostly aware* of her motivations for holding or reaching her interests and preferences, that she is also conscious of the reasons that underwrite them, and that she can give a more-or-less accurate account of them (Christman 1991: 11). These requirements build on the assumption that with a rational agent, there is a connection between one's motivations, her reasons, and, finally, her judgments and the subsequent (collective) decisions to which they contribute. The requirement of having reasons and, in an extended sense, *reasons for one's reasons*, builds off and extends Kant's distinction between *autarchy*, which is the ability of an individual to make her own choices, and *autonomy*, which requires that a chooser is able to state reasons for these choices that reflect their motivations (Elstub 2008). It is the addition of this latter requirement—that one is aware of and able to state their motivations for their reasons—that complicates matters since, as we will see briefly in this chapter and more fully in chapters three and four, motivations can quite quickly become complicated and obscured from an individual's own apprehension (making it complicated to communicate their reasons for having their reasons to another).

The primary goal of this chapter is to defend deliberative democracy as both a fair democratic procedure (again, in the thin sense of it being a procedure that impartial among participants, and a procedure that is acceptable to them) and as an epistemic theory of good and valid democratic judgment and decision making (including the requirement that its participants be autonomous in the extended sense just mentioned). To this end, I will argue that a capacity for autonomy, including the ability to state one's motivations for one's reasons, is the keystone requirement for a theory of deliberative democracy to hold sway as a cogent theory of better democratic decision making. This claim relates specifically to the epistemic

claim made by Warren (2002) that deliberative democracy is more epistemically valid in the context of a liberal democracy (i.e. its decision-making procedures produce better decisions than alternative approaches) than aggregative democracy or other forms of decision making (e.g. elite-driven, technocratic procedures) and also Estlund's (2008) argument that democracy, in general, has more epistemic value *as a procedure* than other forms of decision making.

Recall that I am focusing on autonomy over rationality here since to the extent that autonomy is present, an individual is theoretically capable of giving an accurate and reliable account of *why they have come to some judgment* (i.e. a consistent account that reflects the reasons that *actually motivate* an individual). If that judgment is irrational, it can be interrogated, countered, and corrected. But autonomy is an essential requirement for deliberation since it enables an individual to present accurate reasons for his judgment, which then allows those assembled to engage with it. One could, in theory, give rational explanations that were not accurate accounts of their actual motivations, reasons, etc. Autonomy, when exercised, implies that the account given is a more-or-less accurate one.<sup>14</sup>

The epistemic dimension of deliberative democracy requires a bit more elaboration here, specifically when it comes to the statement that it generates "better" decisions. By "better" decisions, I refer to Warren's claim that deliberation tends to generate decisions based on statements that are "authoritative in relation to [their] referent: facts, norms or subjective desires and experiences" (Warren 2002: 192). This point is essential. In a liberal democracy, asserting an ontological truth is not enough to justify an action because the foundation of

---

<sup>14</sup> In this context I am referring to "accuracy" being related to the specific causes of something, rather than some (conscious or not) rationalization.

democracy is (broadly) built upon the requirement that all those affected by decisions should have some (general) control over how those decisions are made and by whom. Accordingly, Warren's point that deliberation generates better knowledge of a statement's validity (and, subsequently, the reasons for and justifications of a decision) rests on the (accurate) claim that "We gain such knowledge, or confidence in the validity of the statement, through processes of challenge, reason-giving, and verification" (Warren 2002: 192). These processes—undertaken in a deliberation as an intersubjective effort aimed at generating, among other things, shared understanding—are inherent in deliberative democracy. Moreover, they are better achieved and exhibited through deliberation than other means of judgment and decision making, especially in the context of liberal democracy. To put it simply: deliberation fulfills at least three broad essential public goods simultaneously: an epistemic good (i.e. establishing the validity of statements), a political (or democratic) good (i.e. providing grounds for the generation of public and legitimate decisions and including more individuals in decision making), and an ethical good (i.e. recognizing the moral status of those affected by a decision through reason giving and reciprocity).

With this established, in the following pages, drawing on Warren, I explore each of the specific ways deliberative democracy is *epistemically* superior (since this is the focus of my project) to aggregative approaches to democracy to establish deliberation as an approach to decision making more consistent with autonomy. I then present and defend Estlund's (2008) theory of epistemic proceduralism as the strongest and most defensible framework of deliberative democracy as an epistemic theory of collective decision making and judgment, since it provides the scaffolding for understanding how deliberation, autonomy, and

epistemology intersect.<sup>15</sup> I also address Estlund's (2008) critique of epistemic defenses of deliberative democracy: a critique that is grounded in the claim that deliberative democracy cannot generate true beliefs about the common good, since if this is true, then the value of deliberation as an approach to autonomously address moral issues in politics would be undermined. I then examine how autonomy, as it is conceived by Christman (1991), links to deliberative democracy as an essential, keystone capacity for epistemically-valid decision making. I then explain why in order to be autonomous individuals must not be self-deceived. I finish with a short note on why the *epistemological* issue of autonomy and heteronomy in decision making is separate from the broader *ontological* issue of free will and determinism, since the two are likely to be conflated.

## **A note on truth**

One of the most common and famous theories of truth is the "correspondence theory of truth." In general, approaches to truth that fit under the umbrella of correspondence theories hold that a proposition is "true" only insofar as it corresponds to "reality" (whatever that is—that is another debate, and dissertation, altogether). The purpose of this dissertation is not to evaluate the correspondence theory of truth, but rather to make an argument about the capacity for individuals to come together to *establish the validity of statements through a deliberative procedure*. So, my concern here is not whether statements, judgments, and

---

15 While Estlund's framework is concerned with justifying *democracy in general* on epistemic grounds, I adapt it slightly where needed and specify its strengths within the context of *deliberative democracy* in particular. This adaptation is entirely consistent with Estlund's approach, given that the epistemic arguments in favour of democracy in general and democratic deliberation in particular are very similar.

decisions arising from either or both are “true”—that is yet another important question for another project. My concern here is whether or not we can establish a democratic deliberative procedure, series of procedures, or system that can facilitate the establishment of validity through the exchange, by assembled individuals, of preferences and reasons for those preferences. Moreover, I am concerned about how we can generate the necessary good judgments and decisions expected in a liberal democratic society. This is to say that my project is a specific kind of *epistemological* endeavor and not an *ontological* one; while I presuppose that there is indeed both a natural and normative<sup>16</sup> “world out there” that can be apprehended, understood, and shared, my interest is in *how* it is apprehended, understood, and shared, and not in the *content* of what is included in such acts or the physical or metaphysical status of that content.

This interest in and focus on epistemology comes from a more specific interest: how we can collectively make good judgments and decisions. For my purposes, there are two categories of requirements that need to be met for a judgment to be good, each of which is important. First, there is what I am calling the *personal category* in which a “good” judgment is one for which the motivations of the individual making that judgment are *accessible and known*; from which the conclusion follows from the premises (i.e. logically valid); and which we might reasonably expect someone to choose again, affirm, or reaffirm given a similar context. It is “personal” since the focus is on how a given individual produces a judgment. A “good” decision for my purposes is a decision taken collectively—and based on the good judgments of one or

---

<sup>16</sup> By a normative “world out there” I mean a world full of intersubjectively-established moral facts (though not necessarily universal facts) that can be apprehended and discussed.

more of the assembled individuals. Second, there is an *impersonal category* that corresponds to whether a judgment is “correct,” by which I mean valid in regards to the sort of criteria required to evaluate a judgment given the kind of reality it is meant to address: natural, normative, or subjective (corresponding, respectively, to Habermas’ categories of truth, rightness, and truthfulness) (Habermas 1984). It is “impersonal” insofar as its evaluation depends on factors primarily (though not entirely) outside of the individual.<sup>17</sup> My focus here is mostly on good judgments and decisions—I will only address the correctness dimension, as I have defined it, occasionally. It is most important to note that when I do use the term “correct,” it is in this epistemic sense that I mean it, and not some ontological sense.

So, again, my concern in this dissertation is not about how individuals generate “true statements” so much as it is about the pragmatic need to come up with shared epistemic foundations and claims that can be used to reach legitimate, effective, and just democratic outcomes through deliberation.<sup>18</sup> It is also about the ways this process is undermined by common limits to our cognition, and how those are brought about or exacerbated by institutions or common ways of thinking and behaving; just as much, it is about how we might respond to the challenges this presents to democratic life by generating better processes for the ongoing testing of the reasons, judgments, and decisions that come out of democratic deliberation.

---

<sup>17</sup> Specifically, this mode includes whether a statement is valid in relation to the natural world, whether it is right in relation to the intersubjective normative world, and whether it is truthful in relation to others who perceive the speaker as sincere and trustworthy (see Habermas 1984).

<sup>18</sup> My approach here draws on Habermas’ pragmatism (see 1996; 1998); cf. with correspondence theories from Moore (1993 [1901-1902]) and Russell (1994 [1905]).

## Deliberative democracy as epistemology

Warren (2002) argues that deliberative democracy offers a rational way of generating judgments and democratic decisions. He bases this claim on the argument that deliberative processes provide better outcomes than aggregative or other approaches to decision making (rational in this sense refers to a more effective and efficient means to a given ends). In the following section I will present Warren's epistemic argument for the superiority of deliberative democracy over aggregative democracy, and will add two further arguments in support of this claim: First, drawing on Estlund (2008)—that deliberation solves the problem of epistocracy; and second, drawing on Page (2007) and Landemore (2012)—that collective deliberation can address *some* of the shortcomings of bounded rationality that are often present in individual judgment by adjusting deliberative processes to ensure the presence of cognitive diversity (i.e., diversity in perspectives) among participants.<sup>19</sup> Taken together, these arguments help bolster the case for democratic deliberation, as currently conceived, as a way of addressing some challenges to autonomous judgment and decision making. I will revisit the cognitive diversity point again, too, in chapter five.

First, though, let us look at Warren's claim that deliberative democracy generates more rational and valid judgments and decisions. Warren grounds this claim in the supposition that deliberation, insofar as it requires specific procedures, *begins* as a more *epistemically* sound approach to decision making and judgment. To fully explore this point, let us return to and expand a point made above regarding epistemic validity. Warren argues

---

<sup>19</sup> This argument will be revisited in part two of this dissertation when I discuss approaches to designing deliberation and deliberative systems.

Decisions resulting from properly-constituted procedures have more validity—that is, they are likely to be truer or more right or more truthfully related to needs....Validity is, in other words, given by the knowledge that a statement is authoritative in relation to its referent: facts, norms, or subjective desires and experiences. So while a statement may be true (or right or truthful) in some ontological sense, epistemological questions have to do with how we can *know* it to be so. We gain such knowledge, or confidence in the validity of the statement, through processes of challenge, reason-giving, and verification (192).

Thus, the primary function of deliberative democracy as a rational, epistemic theory of democratic judgment and decision making, is the transmission of information from a private, individual state to a public, shared state in which participants can evaluate claims and counter-claims through an intersubjective process that maximizes validity (and good judgments and decisions) through understanding that results from shared reasons. This function implies that there are epistemic norms outside of (and perhaps potentially independent of) ontological “truths” that govern and assist in determining what ends up counting and knowledge. Deliberative democracy as an epistemic theory of democracy draws its rational value from its *reliability* as an epistemic procedure for generating—or tending to generate—*shared knowledge in a given* context; it generates knowledge through a procedure rather than revealing, stipulating, or insisting upon some absolute and non-negotiable ontological reality which may exist in private, but the validity of which, to be shared publicly in a democratic, must

be *defended* rather than *asserted*. This validity-based approach is why deliberation is a more rational way of generating judgments and decisions. As Warren puts it, the legitimacy that follows from deliberation “may become more ‘rational’—not in Max Weber’s sense, which refers rational legitimacy only to the outcomes of positive procedures—but in the sense that legitimacy is generated out of public deliberations which produce reasons with motivational force” (Warren 2002: 187).

It is important to reiterate—as this is the central point of this chapter and, indeed, the animating force behind much of this dissertation—that for deliberative democracy to yield good judgments and decisions, participants must be autonomous in the extended sense of being able to state their motivations for holding their interests, preferences, desires, and, moreover, the reasons for their reasons (i.e. their “actual reasons”). Otherwise, attempts to generate understanding through shared reasons may be undermined. This undermining will occur both within the individual and the group—including attempts at challenge and verification—by either the introduction of false or misleading antecedents or reasons that, quite simply, are not justified given that they lack valid support and are thus suspect. Again, as we will see in chapter three, there are plausible and, indeed, strongly persuasive reasons to question the extent to which agents are capable of acting autonomously in the extended sense, either alone or in groups. Cognitive distortions and limitations are pervasive and are not *easily* eliminated, even at the group level (see Mendelberg 2002, for instance, for a comprehensive review of the ways in which group deliberation can *amplify* these effects). So, there is good reason to ask ourselves whether human beings are well adapted to the epistemic expectations of liberal-democratic citizenship in the deliberative mold (though this is not to say that

improvements are impossible or even unlikely to work, but rather that more empirical work needs to be done *evaluating* deliberation against its normative claims).<sup>20</sup>

## **Politics and expertise in deliberations**

The special concern about *autonomously* (and *rationally*) expressing one's reasoning comes from the fact that the epistemic force of deliberative democracy is, in part, rooted in its purported capacity to act as a theory capable of responding to the twin fundamental requirements of deliberative democratic decision making: (1) that the practice or system establishes or sustains public awareness that there is an issue that we, as a public, need to address, since the outcome will affect us; and (2) that it provides a procedure or series of procedures by which the "we" of the first requirement can collectively decide in a way that is legitimate, procedurally fair, and more likely to be respected by those involved, including those who disagree with the outcome). Each of these imperatives requires shared and common reasons offered by more-or-less equal participants who can express their preferences and the reasons (and reasoning) behind them. Deliberative democracy responds to this need by providing a procedure which makes public the otherwise private reasons, preferences, and motivations of participants, thus generating shared, reliable knowledge than can be used to generate judgments and decisions, provided they are rationally and autonomously generated.

Let me be more precise about this point. Politics emerges at the point at which a decision—broadly conceived—must be taken to respond to or adjudicate among competing

---

<sup>20</sup> This issue of norms and capacities is, ultimately, an empirical question and related to the second half of this dissertation. For a full discussion on this point, see chapters 5-7.

preferences, wills, claims, needs, or desires. This adjudication requires common grounding in basic facts and norms from which the agents or collectives engaged might address the issue at hand through a procedure that is knowledge producing. Ideally, this is established through the force of rational, autonomously-generated statements exchanged between participants. As Warren (2002) argues, in reference to scientific deliberations

[Scientific deliberations] are institutionalized in a positive sense through peer review, discussions among colleagues, scientific meetings, and publications...when others recognize statements as true, they gain in epistemological authority and hence validity. It is important that validity does not come in any direct sense from the ontological 'objectivity' a statement might claim. There is nothing that could establish such objectivity outside of scrutiny of the statement by others... (192-193)

Thus expert claims achieve standing through scrutiny, review, testing (if possible) and, then, if the claim is taken up, public acceptance. The a priori prohibition on merely asserting ontological objectivity (of natural facts, not norms) as a basis for a claim's validity is the condition for that claim to be collectively endorsed. To simplify: it is never enough for an expert judgment to rest on *take my word for it*.

The same can be said about public deliberations concerning political matters, whether carried out by ordinary citizens or their elected representatives. There are strong parallels between scientific deliberation and public, democratic decision making and judgment in

deliberative contexts. Warren (2002) highlights two similarities and one major difference between expert scientific deliberations and public political deliberations, and each reveals something about the epistemic value of deliberation as a political response to divided societies (i.e. most, *likely all*, societies)

First, validity is a product of procedure, suggesting that institutionalized deliberation can establish the epistemic validity of claims and assertions. Second, like science, politics works at the frontiers of validity, although in a different sense. Political issues emerge precisely when epistemic authority is questioned or has yet to be established. But in other ways, politics is distinct from science. In politics, factual issues are intermingled with normative and expressive issues, so that the authority deriving from knowledge of facts is not as easily achieved within political contexts as within the relatively insulated institutions of science...In politics, [experts] must argue and convince. They must enter into deliberations in order to educate, persuade—or be persuaded. Here, expertise has no pre-political rights (193).

Not only should a speaker's claims *not receive* a priori standing in the context of political judgment and decision making, those claims, in such a political context, are a part of a greater endeavour and *must be* evaluated alongside other considerations, including the normative concerns of all those present. In essence, democratic decisions, by their nature, require reasons based on a different, higher standard of admissibility. The process of deliberation, insofar as it

is an epistemic process, is one of generating *publicly valid statements* by communicating private reasons from the individual speaker to the collective, thus subjecting those reasons, much like in the case of scientific deliberations, to scrutiny, review, and, ultimately, to adjudication alongside complementary or competing claims—thus meeting the higher standard required for democratic decisions.

Regardless of the ontological status of any relevant private interest, desire, preference, claim, or experience that an individual may assert as their own, each participant in a deliberation must take part in an epistemic process by which each is given an opportunity—and indeed a responsibility—to establish the validity of that information alongside others through reason giving, verification, and, if necessary, counter-challenge (Warren 2002). Thus, deliberative democracy, insofar as it seeks to produce epistemically valid judgments, offers a reliable, sophisticated, and more rational (in the sense noted above by Warren) way to make decisions collectively by requiring that potential outcomes are subjected to scrutiny and are the result of an exchange of reasons and argument.

### **Two more reasons why deliberation is a more rational way to generate judgments and decisions**

Building on Warren and taking individuals in diverse liberal democracies as the focus of this study, it is relatively easy to further support the claim that deliberation is a more rational approach to decision making by looking at two further benefits it offers, each related to managing political life in diverse communities in different ways. The first relates to making space for and protecting the judgments of the minority and avoiding the problem of epistocracy

(defined below). The second refers to leveraging cognitive diversity for the purpose of generating better decisions and judgments, and enhancing group problem solving. Each of these purported goods gives us reason to support deliberation as a more rational way to make decisions in diverse liberal democracies. Together, if true, they contribute significantly to the appeal of democratic deliberation.

The first additional reason why deliberation is a better way to produce good judgments and decisions (compared to other forms of democratic decision making) is that it, properly designed, solves the challenge of epistocracy. Epistocracy, as defined by Estlund (2008), refers to the rule of experts: it suggests that instead of letting the public (i.e. non-experts) rule, those who 'know best' ought to decide. The problem of epistocracy is grounded in the following logic: If good outcomes are vital to democracy—an issue we will examine more carefully below—and if experts participating in decision making improve the likelihood that a good decision (as defined above) will be reached, then why not let the experts rule? Estlund deftly rejects this proposition on the grounds that a) *authority* does not emerge directly from *expertise* (2008: 4); and, even if we were to stipulate that it did, b) not all citizens in a given polity will agree on *who the experts are* in any given situation and, moreover, not all would necessarily agree on what such experts think should be done (2008: 3-4, 103-105).

Democratic deliberation serves as a rational alternative to rule by experts—though this claim is subject to evaluation depending on the type of deliberative framework advocated (discussed in the next section). Deliberative theories of democracy avoid the problem of epistocracy by taking seriously the claim made by Warren, noted above, that expertise has no pre-political authority, and by grounding the legitimacy of outcomes in a fair and open decision

making process. These theories, when properly constituted, also address the problem of epistocracy by providing a way for those holding minority positions to accept the majority outcome as legitimate without surrendering their judgment (as we will see below).

The second additional way that a deliberative democratic approach to judgment and decision making is more rational than the alternatives is found in a feature of deliberation that enables it to leverage the cognitive diversity of participants to solve problems.<sup>21</sup> Recall that cognitive diversity as defined by Landemore (2013) and Hong and Page (2004) “refers to a diversity of ways of seeing the world, interpreting problems in it, and working out solutions to these problems. It denotes more specifically a diversity of perspectives...interpretations...heuristics...and predictive models” (see also Mill 1999[1859], Page 2007). In deliberative contexts, the presence of a variety of individuals with different ways of thinking has been shown to have the potential to improve reasoning and decision making (Page 2007). Accordingly, it is thus expected to increase the overall quality of deliberation by offering more and different sorts of reasons from which to draw a judgment or decision, and by providing more robust tests for outcomes (Bohman 2006; Druckman 2004; Landemore 2012, 2013; Mercier and Landemore 2012). In this case, the increased rationality of decisions is a product of a practical concern: deliberative democratic procedures marked by a cognitively diverse group of individuals make better, more creative decisions and judgments. Moreover,

---

21 Recall that cognitive diversity does not necessarily require a diversity of individual identities, only a diversity of approaches towards cognition and its related practices (i.e. different ways of thinking). That said, to the extent that differences in identity correlate with cognitive diversity, the two can be thought of as bound up together, though this is not necessarily the case, nor is it logically required.

those outputs are put through more robust tests than they would be if merely left to groups of experts (Page 2007) or, by implication, just put to a vote.

### **Epistemic proceduralism: the best of two worlds?**

Estlund (2008) offers the strongest and most systematic defense of democratic authority as an epistemic theory of decision making and judgment. He calls his approach epistemic proceduralism: a kind of compatibilist theory of democracy, that can be fitted to democratic deliberation, that sits between full proceduralist approaches—generally, those approaches under which the legitimacy of decisions emerges from an *internal* procedural standard, such as fairness in the strict procedural sense of procedural equality—and correctness approaches—under which decisions are legitimate because they are, strictly speaking, correct (e.g. Plato and the rule of the Philosopher Kings; Rousseau and the General Will). Estlund rejects the former on the grounds that democracy is concerned with more than *just* fairness, and, to some extent, is after the ‘right decision’—otherwise, a coin-toss would be, strictly speaking, the most ‘fair’ way to make a decision (since each would have an equal chance of affecting the outcome, which is to say zero chance) and would be sufficient, regardless of outcome (2008: 5-7). He rejects the latter by rejecting epistocracy on the grounds that, as we have just seen, not all in society will agree on *who* the experts are and *what counts* as expert judgment (2008: 3-4, 99, 105). As noted, in so doing, epistemic proceduralism solves the problem of epistocracy by requiring that the minority in a given decision *recognize and accept* the outcome without deferring their judgment to the majority, as correctness theories would have them do (2008: 102-103). This practice makes space for expert judgment and evaluation, among other perspectives, without

privileging experts over citizens or substituting their judgment for the judgment of the participants in any given deliberation. The epistemic force of deliberation, in Estlund's conception, remains in place.

Instead of being rooted in a pure proceduralist or correctness theory, Estlund roots epistemic proceduralism in the claim that properly constituted (deliberative) decision-making procedures are preferable to nondemocratic arrangements, fair (in the sense of being impartial between participants and generally acceptable to a reasonable person), and have a better than random chance of producing the correct decision (i.e. correct by some independent standard, whatever it may be) (Estlund 2008: 17, 98, 102-103, 167-168).<sup>22</sup> According to Estlund, epistemic proceduralism offers "Procedural impartiality among individuals' opinions, but with a tendency to be correct; [it is] the impartial application of intelligence to the moral question at hand" (107-108). So, epistemic proceduralism grounds the legitimacy of its outcomes in a procedure that is both democratic (fair and acceptable) and that has epistemic value—and thus legitimacy—through its tendency to get things right (2008: 105-106). Estlund compares epistemic proceduralism to juries—though not in the sense of Condorcet's Jury Theorem.<sup>23</sup>

---

22 Such a standard could—indeed should—be an item for deliberation, though, as with constitutional matters, one would expect it to be difficult to amend and for such a deliberation to occur rarely.

23 Estlund rejects Condorcet's Jury Theorem—the idea that a sufficiently large assemblage of individuals, each with a more than fifty percent chance of being correct, will reach the correct answer the vast majority of the time since incorrect positions will be washed out by the overall higher probability of correctness. Estlund rejects this on the grounds that: 1) There is no reason to assume that voters are better than random and, more importantly, 2) No one should be forced to substitute the judgment of the majority for their own judgment—especially since one's opposition to the majority judgment gives them further reason to doubt that outcome on top of the dubious likelihood that any given majority is more likely than chance to be correct (Estlund 1997: 185-186). See, related, Page and Shapiro's (1992) use of Condorcet's Theorem to argue that the public *as a public* has stable, rational opinions and refutations by Althaus (1998),

Instead, he argues that juries derive their legitimacy from a fair procedure—an adversarial trial in which both the prosecution and defendant are given resources and a chance to make their case—and their tendency to reach the correct outcome on a better-than-random-chance basis (2008: 8-9, 106-107, 168).

Estlund’s epistemic proceduralism also maintains respect for an ontological reality separate from deliberation without asserting a single, mandatory ‘truth’ to guide all deliberation—all the while making it possible for participants in the system to make ‘correct’ decisions and epistemically valid decisions and judgments. Estlund achieves this by first specifying that epistemic proceduralism is *not* “an ideal epistemic situation” nor is it a “constituting truth” (2008: 19). Instead, Estlund’s framework “imagines deliberators for whom there are independent facts about what ought to be done” (2008: 19). Any slip towards relativism or groundlessness is avoided by claiming that truth, which is “generally unspecific,” is related to some particular framework and thus true or false *within a given moral system* (2008: 5). Thus deliberators can have it both ways: Estlund’s framework recognizes that there is an ontological reality for its participants that they bring with them into deliberations (and perhaps maintained throughout and after deliberations, despite what might be decided within the deliberation). It also accepts that such a reality—with the rights and wrongs and whatever imperatives it might hold—will render specific outcomes within the system right or wrong *for some participants* and not others, but will nonetheless be preferable to alternative decision-making procedures (for the reasons noted above). However, again, the rights of the minority in

---

Kuklinski and Quirk (2000), and Zaller (1992) that are, primarily, grounded in the arguments that voters are unlikely to be correct more often than not and that many of them have a tendency towards *systematic* bias, as I will discuss in chapters three and four.

a given outcome are respected and protected, since minorities are given the option to reject the correctness of the outcome (they need not defer their judgment), even though they are required to recognize and obey it as the product of a legitimate procedure that is fair, and that tends to produce a correct outcome.

Bohman (1998) summarizes the approach this way “...epistemic proceduralism seems to require only that citizens share the assumption that there is some way to judge better or worse reasons and not that a specified theory of public justification provides specific criteria to settle disputes about such norms” (407). This epistemic proceduralism is what Estlund calls a “formal epistemic account”—an account of a decision making that has a “tendency to get things right from the standpoint of justice or common good *whatever the best conception of those might be*” (Estlund 2008: 169, emphasis in original). Democratic deliberation, in this conception of it, thus allows for a political, intersubjective way of working out what ought to be done without compromising the quality of the sorts of decisions reached. However, as I argue below, this requires that individuals exercise their capacity for autonomy.

Going forward, it is important to keep in mind that my approach diverges from Estlund’s in two important ways. First, I rely on a very thin conception of correctness for evaluating democratic judgments or decisions. This conception is based in part on the requirement that validity is included as part of a measure of a good outcome (i.e. validity in regards to the sort of criteria required for evaluating that *kind* of judgment, depending on whether it is about natural, normative, or subjective concerns). Second, while Estlund discusses “independent facts” about “what ought to be done,” I am concerned with *dependent facts* about *what we would like to do*; I suggest that (normative) facts are dependent upon an intersubjective process of acceptance,

just as any decision about what is to be done comes from the authority of a group established to come up with such a decision through the exchange of reasons.

### **Elstub's critique of the epistemic defense of deliberation**

Before turning to a look at how autonomy relates to deliberative democracy as an epistemic approach to democratic decision making, I want to address a critique of the theory's epistemic justification offered by Elstub (2006: 304, 2008: 61-62). Elstub characterizes the epistemic justification of deliberative democracy as being grounded in its claim to offer "the best method of providing good decisions.... because, by generating public reason, it can lead to decisions that are true, well justified or commensurate with justice, needs or the common good" (61). Against this, he claims that it cannot be known whether deliberative approaches to democratic decision making do indeed provide the best method for reaching such decisions since we cannot test this claim; and if we could test it, there could be *another* approach to decision making that is not known, so deliberation cannot be assumed as the best approach. Moreover, he notes, along with Cohen (1989), that there is reason to doubt the existence of a "'real truth' about the common good." Accordingly, he concludes deliberative democracy cannot be justified by reference to its epistemological superiority to other ways of making decisions.

While there is very good reason to doubt that a "real truth about the common good" exists—at least in the absence of a self-contained system in which ethical, social, and political parameters are strictly established and controlled—this fact does not gainsay the epistemic justification of deliberative democracy offered by Estlund (whose full statement of epistemic proceduralism came *after* Elstub's critique). That is because the particular epistemological claim

made by many who justify deliberation on epistemic grounds, and especially the framework offered by Estlund, is not that deliberative democracy reaches or reveals some *ontological truth* about the *content* of the common good; rather, the claim proponents make is that deliberation offers a (fair) procedure by which a given collective can establish validity through reason giving and, if needed, verification and challenge (Warren 2002). They also argue that deliberation tends to generate decisions that are correct (according to Estlund) or good (as I mean it) more often than random *within a given moral, ethical, social, and political system*, while also being based on a democratic (i.e. fair and generally acceptable) procedure (Estlund 2008).

This same point applies to Elstub's critique that there is no test we can apply to deliberative democracy to determine if it does indeed lead to "decisions that are true, justified or commensurate with justice, needs or the common good." Again, deliberation is not about the discovery of ontological truths, but the establishment of epistemic agreement about the validity of statements, and the generation of judgments and decisions about what we want to do in a democracy. Validity in this context is a product of a fair, generally acceptable, and properly constituted procedure. Deliberative democracy, properly structured, offers an approach to establishing valid statements about political issues in the hopes that we might reach some decision in a way that is more rational and thus, hopefully, better than the alternative options and more acceptable to those who will be affected by any decisions that are be made. Deliberative democracy, then, is first and foremost a theory of how a *properly-structured process* can lead to more ethical, rational, and legitimate democratic decisions; it does not *presuppose any particular ethical content* or require reaching common good as an outcome of its proceedings outside of the core requirements and values of liberal democracy

(Warren 2002). Thus, critiques of the epistemic justification of deliberative democracy that draw on doubt about the existence or verifiability of a common good, in the thick ethical sense rather than the thin functional sense, mischaracterize how deliberation is commonly—though perhaps not universally—justified on epistemic grounds. Such a justification is based on democratic deliberation as an inter-subjective process aimed at securing common terms of discussion and evaluation, and facilitating an exchange of reasons to reach an internal understanding of the desires, interests, and preferences of participants—and, sometimes, though not always, a specific political outcome.

### **Autonomy in deliberative democracy: the keystone capacity**

If epistemic proceduralism functions as a framework for the theory and practice of deliberative democracy aimed at generating a foundation for evaluating claims and counter-claims on an ongoing basis, it requires a cognitive underwriter. It is in this sense that I think of autonomy: the capacity that serves to connect the will, through rationality and self-awareness, to the reasons one has for their preferences and judgments in such a way that they can both *reach and understand* such preferences and judgments and their origins; moreover, autonomy requires that individuals be also able to *communicate* that information to others. Autonomy, in the Kantian sense, noted in chapter one, is present when one's thinking is grounded in a sound internal cognitive procedure. Building on this core requirement, I take my conception of autonomy from Christman (1991), who defines it as a state in which "...the influences and conditions that give rise to the desire [or preference or intention] were factors that the agent approved of or did not resist, or would not have resisted had she attended to them, and that

this judgment was or would have been made in a *minimally rational, non-self-deceived manner*” (22, emphasis mine). Thus, in Christman’s formulation, an agent can only be said to be autonomous if she is aware “of the changes and development of her character and of why they came about” (11), since only in this way can she encourage or resist these changes through implied or actual deliberation with others.

There are three reasons why Christman’s understanding of autonomy is appropriate for the study of deliberative democracy that I am undertaking in this dissertation, and I will briefly address each of them here. First, Christman *logically* separates the possibility of a concept of autonomy from that of any concept of freedom, but maintains a link between the two, noting that “...the nature and value of political freedom is intimately connected with the presupposition that actions one is free to do flow from desires and values that are truly an express of the ‘self-government’ of the agent” (Christman 1991: 1). So, at a conceptual level, autonomy, though connected in practice to freedom, is a separate capacity that can be interrogated, understood, and even cultivated, on its own. The concept of autonomy advanced by Christman thus avoids being folded into the concept of freedom as some variation on an absence of restraint. This allows us to discuss autonomy as the internal process of a particular agent—or collection of agents—and provides us with the conceptual opportunity to qualify conditions and expressions of freedom *within* deliberation as being, in part, linked to the extent and quality of the capacities for autonomy held by those participating in the deliberative process.

Second, Christman avoids—deliberately—the tricky problem of infinite regress that is associated with some definitions of autonomy; indeed, he specifically addresses the

problematic definition of autonomy offered by Dworkin (1988) (Christman 1991: 4-10).

Dworkin's formulation of autonomy rests on the claim that a person is autonomous when he is in a position to affirm, reject, or alter his preferences, desires, or goals (Dworkin 1988).

However, as Christman notes, drawing and building on a critique offered by Irving Thalberg, not only does this definition of autonomy open up the possibility that an agent finds himself affirming heteronomous positions (Christman 1991: 5), it invites the problem of infinite regress, since any affirmation or rejection of a preference, desire, or goal will also require its own justification (and affirmation) if it is also to be an autonomous choice. Indeed, any condition of "self-appraisal," as Christman calls it, *vis-à-vis* autonomy, will face this threat.

Christman's definition of autonomy avoids the problem of infinite regress by fixing the condition of autonomy at the *first level* of evaluation: the *process* by which a desire, preference, or interest is *developed* (Christman 1991: 18-19). Nothing would stop an individual from running through this process multiple times—say, to reflect upon a desire that is motivated by another desire; what is essential to the process is the individual's awareness of what contributes to the formation of the preference, interest, desire, and so on.<sup>24</sup> Thus the process of evaluation, if undertaken in conditions of minimal rationality and self-awareness, serves as both the necessary and sufficient condition of autonomy without the need to evaluate any *particular outcome*. The parallel with deliberation rests in the fact the epistemic force of deliberation is itself found in a process, rather than in a moment or an outcome; theories of

---

<sup>24</sup> The standard of what counts as "being aware" is difficult to specify outside of a particular agent and context; however, a minimal, reasonable standard is easy enough to approximate: any motivation that a reasonable individual might realistically be expected to uncover under conditions of self-reflection and having most of the information immediately required in that particular context is enough.

deliberative democracy thus require a concept of autonomy suited to their procedural nature, and Christman's definition fits nicely while also maintaining a critical standard of evaluation.

Finally, and related to the previous point, Christman's conception of autonomy avoids the "time slice" problem of theories of autonomy that ground the capacity for autonomy in the ability to judge a particular outcome at a given moment. Again, when evaluating autonomy in the moment, such as is required by a theory of autonomy such as Dworkin's, there is no guarantee that the agent is not affirming or rejecting or altering a desire, interest, or preference heteronomously. This problem is, once again, avoided by locating autonomy in an *ongoing process of preference/interest/desire formation* and the conditions<sup>25</sup> under which this process unfolds, rather than in a single moment of judgment. The parallel with deliberative democracy in this instance is the same as noted above during the discussion of infinite regress: procedural autonomy is well fitted to deliberative democracy because the efficacy of each as a system is predicated on a process rather than a moment, and, indeed, is agnostic when it comes to the content of a judgment or decision. This principle aligns nicely with the model of deliberation with which I am working here, a model that accepts that an ontological reality for each participant (i.e. beliefs about what is true) will exist in any given deliberation without assuming that everyone shares such a reality and is willing to accept whatever suggested decisions or judgments it implies.

Before continuing to the matter of non-self-deception in autonomy, it will be useful to briefly revisit, summarize, and comment a bit more on Christman's formulation of autonomy,

---

25 It is worth mentioning that as Christman notes "...these conditions may have little to do with how the agent evaluates the desire [or interest or preference] itself..." (1991: 11).

since it runs counter to most approaches to defining the concept. According to him, the conditions for autonomy are

- (i) A person P is autonomous relative to some desire D if it is the case that P did not resist the development of D when attending to this process of development, or P *would not have* resisted that development had P attended to the process;
- (ii) The lack of resistance to the development of D did not take place (or would not have) under the influence of factors that inhibit self-reflection;  
  
and
- (iii) The self-reflection involved in condition (i) is (minimally) rational and involves no self-deception (Christman 1991: 11, emphasis in original).

The meanings of (i) and (ii) are each reasonably self-evident and have, anyway, been discussed above. However, there are two elements of Christman's definition of autonomy upon which I wish to dwell, since they are directly relevant, though perhaps not obviously so, to an understanding of how and why autonomy is necessary for deliberative democracy to be plausibly considered as a theory of democratic engagement aimed at generating epistemically valid outcomes.

First, autonomy requires what Christman calls "minimal 'internal' conditions for rationality" (14). He cites a basic consistency of beliefs and desires as requirements but stops short of demanding that there be an absolute and clear link between the epistemic process of

developing internal consistency and the ontological objectivity of the external world. Thus autonomy requires only *internal* consistency, and not an objectively verifiable connection to some pre-established ontological reality. The link to the model of deliberative democracy that I am working with, and to its epistemic defense, is clear enough: participants in deliberation who are in search of epistemic validity must be capable of maintaining at least a basic internal consistency, otherwise the grounding upon which the deliberative enterprise rests is subject to its own inconsistency (though this does not necessarily require universal agreement on any given ontological reality). *Epistemically*, deliberation is primarily about process and not outcome, though proponents make claims about the sorts of outcomes that are likely to come about from well-designed and executed deliberations; it is not enough to come to ‘the right’ decision, since it is not clear on what bases we would evaluate this absent some justifying procedure—which is what deliberative democracy is: a way of doing things. That decision must also reflect a logical consistency that is unlikely to emerge if it is drawn from a collection of illogical internal processes; and even if it did, it could not be said to be the product of an epistemically-valid process.

Second, autonomy requires that “the influences and conditions” surrounding a judgment, through the interests, preferences, motivations, and desires that support such a judgment, were approved of by the agent—or would have been—under what we might call conditions of sufficient awareness (a minimal level of knowledge about factors relevant to the judgment at hand, which I will discuss below). In other words, the agent must be in a position to assent to all of the immediate factors that contribute to their judgment. This assent requires, as Christman notes (11), self-reflection and, I add, returning to the extended conception of Kant’s

autonomy over autarchy, it also requires that the factors considered be *the actual mobilizing agents related to the judgment* (or desire, preference, etc.). This is to say that an agent is only autonomous to the extent that the process of self-reflection he undertakes in the course of approving of a judgment *accurately* links “influences and conditions” to outcomes and is not interrupted by some internal or external force. This requirement of autonomy is what I am referring to as the principle of non-self-deception (examined further below).

Violations of this principle are common. In chapters three and four I will deal with these violations in much greater detail by looking at four cognitive models that explain how certain modes of cognition explain why cognitive distortion is so common; but, for now, one key study from political behaviour research will serve to illustrate the form of the problem. Sniderman et al. (1986) set out to understand how citizens come to decide what they think about some particular complex political issue. They found that among *both* more and less educated voters, broken reasoning chains were common: rather than building from a premise or a series of premises to a conclusion through a series of inductive steps, many respondents began with a preference and *worked backward to rationalize it outside of their awareness*. Low-information respondents drew on *affect-driven* considerations, while high-information voters drew on *ideology*—but each *rationalized* their judgments. While neither of these processes of chain reasoning may be particularly devastating to autonomy in opinion formation on their own, Sniderman and his colleagues found that, in general, individuals in mass publics maintain consistency of belief by adjusting their policy preferences to their likes and dislikes, and then by *reasoning backward from their conclusion*. Thus, for some citizens—at least on some issues—preference formation is about the ex post facto rationalization of prior affective or ideological

relationships (for more expressly political examples of this, see Lodge et al. 1995; Lodge and Taber 2013). This finding implies that whatever the *conscious* approach that some take to explain their preferences or beliefs, knowledge of *how* they were formed and *what drives them* may remain elusive. Given some respondents' rationalizations, they, strictly speaking, cannot be said to be behaving autonomously, since the link between influences and conditions and outcomes is broken or distorted. This phenomenon is an example of an agent expressing autarchy but not autonomy since the agent might not have approved of the outcome if they had the full appreciation of their motivations and better knowledge of the process by which they came to their judgment.<sup>26</sup>

In the case of deliberative democracy and the procedural generation of epistemic validity, the same standard can be upheld: only reasons given that *accurately* reflect influences and conditions can count as being part of an autonomous judgment. Establishing epistemic validity requires not only an exchange of reasons directed at establishing common terms of discourse, but also a common working understanding of the interests, desires, motivations, preferences (and so on), and the subsequent reasons that support them. Moreover, it requires, at least, a working agreement on basic facts and norms, and that each reason is logically and accurately attached to its motivational source. If parties in a deliberation are making judgments that are disconnected or hidden from their ultimate motivations, they cannot be said to be behaving autonomously—which undermines any claim to be producing epistemically valid

---

26 See also the related phenomenon of “motivated reasoning,” which refers a process by which issue partisans process a variety of information, some of it challenging to their position, in a way that confirms or even enhances their pre-existing position (Ditto and Lopez 1992; Taber et al. 2001). The phenomenon of motivated reasoning will be thoroughly explored in chapter four.

outcomes, and thus calls into question what deliberative democracy really requires as sufficient justification for claims made to one another by participants. Why? Because reasons that are partial, distorted, or which rely on hidden motivations prevent potentially relevant information from being put on the table during a deliberation, thus constraining the ability of those assembled to establish the sort of validity required to make good judgments and decisions.

So, for it to be said that an agent is behaving autonomously during deliberation, she must be able to maintain *basic and minimal* internal logical consistency and, upon reflection, be able to more-or-less accurately link her motivations to the reasons she is offering towards some particular outcome: *if too much is lost in translation between one's motivation and one's stated preferences and the reasons given for those preferences, the agent cannot be acting autonomously*. Moreover, for the reasons noted above, for it to be said that a deliberation is working towards establishing epistemic validity, its participants, at the very least those who are contributing significantly to the outcome of the deliberations, must meet the basic requirements of autonomy.

### **Autonomy and the principle of non-self-deception**

I have already discussed, via Christman's definition of the concept, what I mean by autonomy and I have laid out its core requirement of non-self-deception: the requirement that one's motivations are reasonably aware to them and that the reasons one gives for a judgment are those which *actually motivate them*. However, it remains to be seen *why* this principle is so essential to autonomy. I have already addressed why Dworkin's formulation of autonomy as a capacity to affirm or reject outcomes after the fact invites the problem of infinite regress. Also,

and again, as this point is essential, his definition invites the possibility that one is affirming heteronomous sources without awareness of what these sources are or how they impact one's decisions or judgments (a corollary of the infinite regress problem). Autonomy requires, for the reasons noted above, non-self-deception to avoid a state in which one's motivations and consequent preferences or judgments are determined outside of their awareness.

Nonetheless, neither of these counter-arguments explains *why* non-self-deception is important to judgments and decisions in general or in instances of deliberation in particular. If, for example, a deliberative procedure is fair and acceptable, and has the tendency to produce good outcomes, why, then, is it *also* important that participants in the system not be self-deceived? Why is it important that they not only produce often-enough good (i.e. more often than random) and legitimate decisions but *also* act autonomously? We might call this the problem of *a rose by any other name*, since it poses the challenge of whether some ersatz copy of a deliberative system and its participants, though of different fundamental composition (i.e. without autonomous agents), is *a priori undesirable* and less of an appropriate standard for judgment and decision making in liberal democracies.

There are at least two plausible lines of argument against the potential banality and irrelevance of self-deception and heteronomy to democratic deliberation—arguments against the ersatz copy. The first is empirical: If self-deception and heteronomy and their effects were randomly distributed both among and within participants in a deliberation, then the concern about those effects might be somewhat mitigated or perhaps even eliminated altogether. However, those effects *are not always randomly distributed*. Part of the problem of self-

deception and heteronomy is that their presence and effects can become *structurally distributed* in at least two ways.

The first way concerns *among whom* their effects are most pronounced: ample research shows that low-information individuals (i.e. those who may be engaged with politics but who are poorly or partially informed) are more affected by heuristics whose true effect they are unlikely to be aware of—and so the system, as-is, tends towards structural bias in the outcomes it produces (see, for instance, Althaus 1998, Chong and Druckman 2007, Converse 1964, Cutler 2002, Iyengar et al. 1982, Kahneman 2011, Kuklinski and Quirk 2000, Sniderman et al. 1986, Zaller 1992; cf. Lenz 2009). The second concerns *who tends to bear the burden of the adverse outcomes* of heteronymous judgments or preferences (however they might be generated): for example, people of colour (Mendelberg 2001, Sniderman et al. 1986), the poor and undereducated (Althaus 1998, Frank 2004, Zaller 1992), and other groups who already suffer the effects of negative stereotyping (Kuklinski and Quirk 2000).

Whatever the many the complex reasons why the presence and effects of self-deception and heteronomy are distributed, the deleterious effects of their presence on deliberations and the outcomes of deliberations are significant, especially among certain populations of (often) already disadvantaged groups, while the benefits of potential positive effects tend to be accumulated by those who are already advantaged by the status quo or who begin from positions of relative power or influence. For instance, the threat and effects of climate change—e.g. heat waves and extreme weather events—are disproportionately borne by the poor (United States Environmental Protection Agency 2015). In an experiment on motivated reasoning (discussed further in chapter 4) and preferences on climate change policy in the

United States, Hart and Nisbet (2012) find that political partisanship influences support for climate change and that new information—shared equally and presented identically to each participant—*further polarizes* opinions on climate change between Republicans and Democrats (see also Taber, Cann, and Kucsova 2009). And if such effects are combined with a group in which a minority-type is outnumbered, polarization can become worse through increased (non-cognitive) bias and decreased cooperation (Bettencourt and Dorr 1998). So, in the case of climate change, to the extent that polarization leads to inaction or inadequate action—which, to date, globally, it largely has—then cognitive distortion (which undermines autonomy) contributes to the structural and persistent disadvantage of certain groups over others.

Even if the effects of self-deception and heteronomy were randomly distributed, there is no reason to believe that they would contribute to good decisions on a better-than-random basis, while there are many reasons to believe that autonomy and non-self-deception would. Imagine trying to assemble a piece of furniture. The heteronomous/self-deception approach is akin to winging it: grabbing the hex key and slapping pieces together under the impression that “Hey, I know what I’m doing here.” The autonomous approach would be that which requires that one had a reason to connect particular bits and pieces—this leg to that frame—and, moreover, reasons for those reasons—to stabilize the piece, to keep it from falling over, etc. The specific analogy here, of course, related to approaches that link epistemically valid claims and reasons to judgments or decisions and those that do not. Being autonomous, having reasons and knowing *why* one holds them, should increase the probability of good (or correct) decisions by increasing the reliability of both the information one has and the process by which

one internally evaluates that information, connects it to reasons, and reaches purposive and considered judgments.

The second argument against heteronomy and self-deception is theoretical and aims to respond to the hardest test of the value of autonomy and epistemic validity. Imagine that the presence and effects of self-deception and heteronomy were both randomly distributed among participants in a deliberation, that there was no structural bias against any group, and *still*, the outcome of the process was guaranteed to have a better-than-random probability of reaching a good result. What, then, would be the problem with self-deception and heteronomy? The empirical concerns raised above would no longer apply since the process would be guaranteed to be, at least, equal to a process that was fair, supportive of autonomy, and included a tendency to produce correct, just, and epistemically valid outcomes.

In this (extremely unlikely) instance, the value of autonomy becomes separated from any particular outcome and, indeed, leaves the realm of empirical concerns (almost) all together since we have stipulated that agents in this hypothetical (and, again, highly unlikely) scenario marked by heteronomy and self-deception would be functionally as sound as those who were autonomous and fully aware of their reasons and their motivations for having those preferences. Here we need to consider the *inherent value of self-determination*. A significant part of the value of autonomy and non-self-deception is derived from the tendency we have as humans to *desire* self-determination as an organizing principle for our lives, both individually and collectively, separate from whatever other functions it serves. To the extent that one is heteronomously directed and self-deceived, she cannot be said to be self-determining, since, by definition, heteronomy implies the determining work of an outside force; therefore, self-

deception is logically inconsistent with self-determination, which, by definition, requires self-awareness as well as the opportunity to put into practices one's desires, goals, and so forth.

Of course, complete autonomy—what is known as maximal autonomy and is defined as radical, boundless, and fully-independent self-creation (Berofsky 1995)—is an impossible and even counter-productive standard: the capacity to define and develop oneself outside of *any* external determining forces is an illusion, and it is not clear that it would be desirable if it were even possible. It does not take much to demonstrate how such an approach to understanding autonomy quickly becomes incoherent: after all, we are born into a time and place, each of which has its own social, political, and technological context that begin to shape us from the moment we are born—and in some ways, even before that. We are also born with unique bodies prone to different affective states and imbued with diverse mental and physical abilities. These realities conspire with others—local and global, cultural and biological—to generate broad paths for us to trek as we develop. We cannot choose to live in a neutral environment, nor can we choose neutral minds or bodies, and so the concept of maximal autonomy fails to allow for a definition of autonomy that preserves any hope of reaching a functional level of behavioural correspondence to the definition.

However, returning to Christman's more constrained definition of autonomy as the rational, non-self-deceptive, acceptance of the conditions of self-development and choice, we can see that this functional understanding autonomy allows for self-determination while also respecting that some external determination is necessary. *It is the process of autonomous thought, the capacity to rationally explain and reflexively accept or deny our preferences or judgments* that makes us autonomous and, thus, self-determining. Without this capacity, one of

the important (though neither necessary nor sufficient) elements of what makes us human—and deserving of/capable of engaging with and maintaining liberal democratic governments—is fundamentally undermined.

So, an absence of self-deception and heteronomy, and by implication, the presence of autonomy, is necessary for deliberation—and elsewhere—for (at least) two sets of reasons. First, because we have good reason to believe that autonomous decisions are more likely to be epistemically valid and well-reasoned than those based on self-deception and heteronomy, and because they are also less likely to include structural bias both among specific groups of deliberators (who suffers from structural bias) and among those who are affected by decisions conditioned by such bias (whom such suffering affects through decisions that are made by the group in question). And, second, self-deception and heteronomy threaten self-determination and what individuals tend to want for themselves: the ongoing opportunity to remain apprised of the relevant information that has some a bearing on that process of self-determination. However, as we will see in chapters three and four, structural cognitive distortions—which are common and pervasive—are a threat to autonomy and the power of deliberation to generate correct, epistemically valid judgments and decisions.

### **Self-determination and democracy**

Having discussed self-determination as it relates to the individual and how he directs his life, it is worth saying a little bit about why theorists of democracy in general and deliberative democracy in particular should care about self-determination insofar as it related to *democratic* self-determination collectively. While the principle is often taken for granted as important and

desirable for a democracy, that does not exempt us from the need to understand *why and to what end*. In the context of democratic theory, self-determination as a guiding principle recognizes *the inherent tendency for individuals who must live together in a bounded space to be different from one another and to disagree*. Even individuals raised in homogenous communities show variation and unique identities, and they come into conflict with one another. Self-determination thus scales nicely from the level of the individual who can and does—and usually seems to want to—be the author of their life to that of the community (or city, province, or country) that wants to do the same. Historically, with few, if any, enduring exceptions, individuals and communities have tended towards self-determination—though to varying degrees of success and often only after much (indeed, ongoing) struggle.

As sceptical as we should be about claims about “human nature,” one thing is certain—when given a chance, human beings tend to strive for individual or collective (or both) self-determination, however inter-subjectively bounded that self-determination may be. Moreover, at least in the modern age, democracy has often been the superior mode of organization for enabling broad self-determination among groups of people with persistent and divergent interests and preferences (though not universally so); and within the context of contemporary democracy, the (individual) capacity for self-determination (enabled in part by the capacity for autonomy) enables democratic citizenship by giving individuals a chance to be active participants in self-government rather than passive objects of governance. Deliberative democracy as a form of democratic practice in decision making is particularly well-suited to self-determination because it gives individuals—whether citizens or their elected representatives—a chance to substantively participate in exchanging reasons for and against a

range of preferences, and to use their capacities for autonomy to generate judgments in the pursuit of self-determination. Also, while allowing individuals to practice self-determination through reason-giving, deliberation also requires that participants *recognize one another as participants*, thus setting up a relationship in which they are both recognizers of and recognized as active citizens in the pursuit of self-determination at both the individual and collective levels. Through this process, individuals are not only recognized *as citizens*, but also, as Arendt or Kant might suggest, as human beings.

The reach of the principle of self-determination is potentially quite far. As Nedelsky and Beiner put it “If it can be shown...that the quality of our experience atrophies in proportion as we passively yield to the judgments of others and cede greater and greater dimensions of political responsibility, then we would have powerful reasons to believe that active citizenship is a major component of the human good” (2001: ix). While it is outside of the scope of this dissertation to fully pursue that “if,” it seems intuitively true that this is indeed the case. Moreover, as noted above, human history seems to bear out this point in many, if perhaps not all, cases. To Nedelsky and Beiner’s point I would add that active citizenship is not just a component of citizenship or the human good, but an important—though not necessary—component of human life as such, at least insofar as it creates space for pursuing and realizing self-determination.<sup>27</sup>

---

<sup>27</sup> In chapter five in a section entitled “*Tradeoffs and limits: Who can and who cannot deliberate*” I note that a focus on democratic deliberation in general and on autonomous deliberation as I have conceived of it in particular means that certain individuals may be excluded from my analysis (e.g. due to severe cognitive disability). In that section I argue why

## Relational autonomy

Now, returning to autonomy, it is, I have argued, a *capacity exercised through a process* (Nedelsky 2001, 2012). A person who is exercising their capacity for autonomy can account for their motivations and can give valid reasons for their choices when reaching a judgment (or taking action). To the extent that they are unable to do so—either because they cannot give *any reason or because the reasons they give are not the reflections of actual motivations or reasons*—they are less autonomous and less self-determining. But, so far my discussion of autonomy has been rooted in individual behaviour and cognition; and yet, as I have begun to show in this chapter, and will elaborate upon in chapters three and four, so much of human behaviour and cognition is *socially determined*. How, then, can an understanding of individual autonomy as I conceive of it be reconciled with the reality of human life as the product of a myriad of day-to-day relationships carried out over a lifetime?

The work of Jennifer Nedelsky on relational autonomy helps bridge this gap. As she argues, “What makes autonomy possible is not being independent of all others, but constructive relationships—with parents, teachers, friends, colleagues and officials of the state. Autonomy is thus also not a characteristic that we simply achieve...its flourishing depends on the kinds of relationships...of which we are a part” (Nedelsky 2001: 111; 2012). She lists “biases, fears, emotions that cloud rather than facilitate judgment” (111) as challenges that emerge from a focus on “private considerations.” Indeed, she argues, echoing Kant, that to the extent that these or similar factors drive our “judgments,” we are not making judgments at all,

---

this fact, while being far from ideal, does not present a significant challenge to what I am attempting to argue in this dissertation.

but have, rather, a certain “emotion or conception” (111) *masquerading as a judgment*. She prescribes taking multiple perspectives when forming judgments to combat the deleterious and autonomy-reducing effects (including the various dimensions *of the self*).

Nedelsky’s conception of autonomy as a relationally-bound capacity is complementary to and consistent with the conception of autonomy that I use in this dissertation. While my conception of autonomy, taken from Christman, is an individually-executed process of judgment formation and affirmation/review, Nedelsky’s relational understanding of autonomy both complements and contextualizes the one I use. Her understanding of relational autonomy is based on autonomy being *enabled by* social relations; still, it complements Christman’s conception of autonomy by providing a perspective that can be taken within the process of reviewing one’s preferences. For instance, when I review whether some particular judgment is (truly) autonomously reached, I can employ multiple perspectives as tests designed to interrogate my motivations. Indeed, as I will argue in chapter five, there are specific tactics that can be used to do just this within a deliberative setting. Moreover, Nedelsky’s understanding of autonomy as relational also contextualizes my understanding of autonomy as an individual process of self-reflection and/or procedural checks and balances during the process of forming a judgment. Specifically, Nedelsky characterizes autonomy as inherently bound up in relationships with others—so that the individual is socially-bound and constituted (2012). Accordingly, her perspective provides another argument against the maximal-autonomy perspective of radical, unbounded self-determination while revealing tactics for enhancing an *individual* capacity for autonomy.

That said, there is a difference between these two conceptions of autonomy—they are not wholly interchangeable. Ultimately, for this dissertation, I think of autonomy as a process conditioned collectively but expressed individually. Its parameters are set externally and intersubjectively, but it is the individual who must act autonomously or not, which may not always require others. Granted, by taking the perspectives of others, it may help the individual to uncover their motivations; but, in the end, autonomy requires that it be the individual who is practicing their capacity for autonomy (or reaching it more or less relative to a continuum—which is based on an approach for measuring autonomy that I will outline in chapter five) or not. So, while autonomy may be relational in terms of its broad constitution, the capacity itself is reached and expressed individually—which is my concern in this dissertation. I do not think this limit is inconsistent with Nedelsky’s conception of autonomy; indeed, if anything, each of the two conceptions makes the other more complete and useful for understanding how human beings reach better or worse judgments.

### **A note on the philosophical debate between free will and determinism**

In closing this chapter, I want to note that a discussion of epistemic autonomy and heteronomy in deliberative contexts, specifically as each relates to judgments and decisions, is *adjacent to* the broader discussion of the *causal properties* of free will versus determinism (not self-determination) in an *ontological sense*. As cognitive neuroscientist Michael Gazzaniga defines it, “...determinism is the philosophical belief that all current and future events, actions, including human cognition, decisions, and behavior are *causally* necessitated by preceding events combined with the laws of nature. The corollary, then, is that every event, action, et cetera, is

predetermined and can in principle be predicted in advance, if all parameters are known” (Gazzaniga 2011: 111).

Whatever the merits or demerits of either camp’s position concerning free will and determinism, the question of what sorts of immediate cognitive factors go into producing an individual judgment or collective decision can be sectioned off as a separate issue, perfectly consistent with either perspective, and interrogated. That is because, as I have briefly touched on, and as I will discuss in greater depth in chapters five to seven, decisions can be made to come from more or from less autonomous agents, or from more or less self-deceived agents, depending on factors such as education, self-awareness, attention to detail, motivation to reflect and engage in thought, the structure and composition of deliberations, and so forth. Whatever the macro-antecedent causes of the states in which judgments or decisions are made, those within those states, those making decisions and judgments, can be reliably evaluated as being *more or less autonomous* according to a prescribed standard; and, moreover, they can be *made* more or less autonomous (as I will argue in chapter five) through changes to personal practices and institutional arrangements. Thus, the level of analysis with which this dissertation is concerned—the individual as a knowing and thinking agent in a deliberative context—is of interest independently of any broader ontological questions about free will and determinism as discussed above. To the extent that I am interested in freedom, it is of a variety very similar to that conceived by Kant—the replacement of non-rational causes of behaviour (e.g. external forces, subterranean desires) with causal reasons, though in my conception I make more room for “desires” and other affective considerations as legitimate

grounds for reasons, so long as an individual can communicate their *actual reasons* for reaching a judgment or for taking an action.

## **Conclusion and summary**

In this chapter, I argued that deliberative democracy is a theory of how ordinary citizens can, among other things, make epistemically good decisions in a fair way (i.e. a way that is impartial among participants and generally acceptable to all who engage in democratic deliberation) on a better-than-chance basis. I also argued that most of theories of deliberative democracy require autonomous agents to cash out these promises, but that autonomy stands out as the “keystone capacity” required for good democratic deliberation. I have used Christman’s definition of autonomy as the capacity of an agent to generate judgments that he or she did not or would not have resisted, in a theoretically self-reflective, non-self-deceptive, and rational way, which I think is the most appropriate conception of autonomy for this dissertation. Furthermore, I explored Estlund’s argument that deliberative democracy is defensible on the grounds that it does indeed fulfill this epistemic process through “epistemic proceduralism”; in the process, I have also distinguished, using the work of Estlund and Warren, between expertise in deliberation and ordinary citizen deliberation. I also compared my conception of autonomy with the (complementary) conception offered by Nedelsky—relational autonomy. Finally, I distinguished autonomy (and self-determination) from free-will and determinism, noting that I am interested here in the former and not the latter.

## Chapter 3: Challenges to autonomy in democratic deliberation

The problem of autonomy, as it relates to deliberative democracy as an epistemic theory of better judgment and democratic decision making, arises from a more fundamental problem that runs through many, if not most, of our social, political, and economic systems: there is a significant gap between our nature and our culture—between what our capacities tend to allow us to accomplish, and what is normatively expected of us. This claim, when made in a constrained sense, is both intuitively appealing and empirically demonstrable, but is often treated as more controversial than it ought to be. Such controversy is mostly misguided, at least when the nature/culture gap argument appears in the more constrained sense by which I mean it in this dissertation. Much of the controversy surrounding this gap, I believe, stems from poorly defined or underdefined terminology used by those who make this argument, or because definitions of either nature or culture that tend to either overreach or over-specify what they can reliably and empirically demonstrate. I will try to avoid this common problem while I characterize this gap before moving on to discuss precisely how it impacts the possibility of autonomous deliberation in contemporary liberal democracies.

When I say that there is a gap between nature and culture, I simply mean that when it comes to certain behaviour, practices, norms, and imperatives, such as democratic deliberation, what is *generally expected of us* and *what we can reasonably be expected to deliver* tend to regularly and systematically diverge. This gap emerges because the expectations and standards we adopt often outpace our ability to meet them. On balance, this is an effect of our bounded rationality and the significant impact of a-rational modes of cognition on our day-

to-day lives (Heath 2014). By “nature” I am referring simply to the constrained evolutionary cognitive capacities that most human beings have developed and the behavioural tendencies—extremely well documented by social and cognitive psychology, sociology, political psychology, economics, and other fields—that *tend to be* concomitant with those capacities. I do not mean to suggest that human nature is universal and absolute or eternally fixed or fully-determining; instead, I mean to suggest that our nature is, fundamentally, a condition that *limits the range of behaviours* that are possible for us, sets *certain cognitive limitations* that may exist prior to our expectations of behaviour, and that *tends to yield*, on balance, certain predictable behavioural patterns or responses in the presence of specific stimuli or contexts, regardless of what our cultural expectations might be.

By “culture” and institutional arrangements I am referring broadly to the malleable (though they tend to change only with time and effort) *time and space specific* sets of practices, rules, norms, expectations, and standards that emerge from, among other things, evolutionary imperatives, biological necessities, historical particularities, ideational supporting structures, and material circumstances. Whatever the origin of expectations that emerge from these structures, there is no *a priori* guarantee that such expectations will align with our natural abilities in every case or instance since *what we imagine or expect* human beings to be capable is not necessarily the same as *what we are actually capable of*. This is the nature/culture divide that frames and informs this chapter, and indeed much of this dissertation.

Regardless of whatever causal mechanisms might be at work in producing these circumstances, regardless of the balance of causal force between idealism and materialism, free will and determinism, and other macro-causal theories, it is indisputable, and unremarkable

that the possibilities and limitations of our minds have an impact on our behaviour, including our expectations. The specific case with which I am concerned in this dissertation is *the potential normative overreach by theorists of deliberation who implicitly or explicitly rely on an account of human agency and cognition that tends not to align with agents typically deliver*. I am equally concerned with the question of what we might do about this in order to contribute to generating better political judgments and decisions—that is, judgments and decisions that come closer to meeting the high standards and expectations we set for ourselves.

In this chapter, to demonstrate how this overreach affects deliberative democracy, I will do four things. First, I will outline what theorists of deliberative democracy expect from deliberative agents, specifically focusing on the normative content of those expectations—i.e. what is desired of deliberation and those who deliberate. Second, I will summarize and analyze a broad theory of cognition drawn from social and evolutionary psychology—systems theory—which divides cognition into two streams (known as a dual-process theory). The fields of social and evolutionary psychology are concerned with, among other things, showing that certain cognitive limitations are built into the structures of human brains and are expressed as thought patterns while tending to present themselves in certain behavioural patterns. This phenomenon is perhaps, though not always, influenced by culture in certain expressions, but generally independent of it. Looking at dual process theories in general, it becomes apparent that certain challenges are implicit in many cognitive undertakings, including democratic participation in general and deliberation in particular. Third, I will lay out four leading models of cognition from social psychology that demonstrate some of the specific behavioural implications of social and evolutionary psychology in the context of contemporary societies in

general and, more specifically, in deliberative contexts (I will tie them back to the systems theory when necessary). The first model is Richard E. Petty and John T. Cacioppo's elaboration likelihood model (ELM); the second is John Jost's system justification model; the third is Jonathan Haidt's account of moral intuitionism, and the final is John Bargh and Tanya Chartrand's automaticity model. Finally, I will conclude by summarizing the implications for deliberative democracy that are generated by, in general, our evolutionary cognitive capacities and, specifically, by the four models examined in this chapter and suggest how theorists and practitioners of deliberation might respond to the challenges that emerge from those implications.

While these challenges to autonomous deliberation and good judgment and decision making are significant and entrenched, I think they can be addressed in productive ways. Ultimately, this dissertation is about highlighting cognitive challenges to deliberation revealed by research in political and social psychology; it is also about developing, *conceptually*, personal practices and approaches to institutional design to address those challenges head on. In that way, I think the gap between our nature and our culture—at least when it comes to democratic deliberation—can be bridged in a way that will produce better judgments and decisions.

### **A note on epistemology, the brain, and our environment**

This chapter is rooted in an epistemological understanding of the human being and is not concerned with ontological explorations. I am interested in how we come to know what we know, both individually and collectively, and how that affects collective knowledge, judgment, and decision making. Because this *epistemological* approach is so central to this dissertation, it

is worth saying a bit more about it. To do so, in this section, I will draw on the work of biologist Gerald Edelman to discuss this approach that underlines this project. More precisely, I take as my point of departure his concept of “brain-based epistemology” (Edelman 2006). According to Edelman, this approach “...refers to efforts to ground the theory of knowledge in an understanding of how the brain works” (Edelman 2006: 2). His interest in epistemology and cognition is largely, unlike my own, concerned with consciousness. However, insofar as he is concerned with how epistemology relates to cognition within a triad that includes brain, body, and environment, his approach to understanding how we come to know things, given the types of brains we have evolved, is useful for the matter at hand. Indeed, such an approach both allows for a broader enquiry into knowing that allows space for examining ideational, biological, and environmental factors, and we can place it squarely and soundly within the traditions of both evolutionary psychology and evolutionary biology, which are useful for understanding how (and why) our nature and culture might have split in certain areas. This is important to keep in mind in reading this chapter given that the underlying theory of epistemology that supports my arguments is one concerned with an interaction between brain/psychology, bodily processes, and the environment: the kind of general context that deliberators will tend to find themselves in when faced with generating political judgments and decisions. If we can understand our thinking and reasoning as guided by several sorts of conditions, then we can begin to unravel how to best generate knowledge, share it, reach our own judgments, and come to share those judgments with others to make decisions—which is what democratic deliberation is all about.

## What is expected and required of deliberative agents?

Theories and theorists of deliberative democracy tend to both assume and require autonomous and rational agents capable of exchanging reasons for their political preferences. Participants in a deliberation are also expected to be open to being swayed by the force of the better argument (as Habermas puts it). In general, when we consider deliberation as a better—or the best—way of making decisions, some combination of democratic goods and normative preferences are included as justification for this claim—for instance, a better-than-random chance at producing good decisions, the increased political legitimacy of those decisions, respect for the agency of individuals, the minimization of the potential for violence in politics, and so on. These goods and desires are what I am calling the normative expectations of deliberation, since these are outcomes or justifications that are tied to a particular place, time, and theoretical program: namely, a philosophical tradition with its origins in the European Enlightenment and developed in the context of twenty-first-century liberal democracy. As noted above, the very fact that such expectations are *normative* in a broad sense means that they are artifacts subject to limits, and there is no *a priori guarantee* that their realization is possible, probable, or even, ultimately desirable given what I am calling natural constraints—constraints based on cognitive or other biological particularities. In the case of this dissertation and the argument I am making, the context is twenty-first-century liberal democracies marked by diversity, pervasive disagreement, speed of communication, the complexity of our social and political systems, and the status of certain capacities (e.g. rationality and autonomy).

Regardless of whatever challenges are posed to theories of deliberation by cognitive limitations, the desired goods themselves can be evaluated amongst one another against a

standard of internal coherence—whether the requirement of each is consistent with the presence of the others. For many—likely most—theorists of deliberation, at least four general categories of democratic goods and normative standard can be discerned from their theories of deliberative democracy: the expression of rational judgments and decisions, the provision of space for the practice of self-determination and autonomy, the generation of valid reasons for political judgments and decisions, and the production of more legitimate political outcomes. These can be further broken down into sub-categories, with the former two—rationality and self determination/autonomy—relating to specific capacities of *deliberators*, and the latter two relating to functions of *the system of deliberation*. All four taken together produce an argument about the value of deliberation as such.

### **The expression of rationality and self-determination/autonomy**

Let us look first at each of the two individual-level assumptions and requirements of theories of deliberation: rationality and self-determination/autonomy. Each requirement implies assumptions about the cognitive capacities of individuals who deliberate. Theories of deliberation assume rational agents, though this assumption unless very narrowly and unfairly construed, does not imply that agents follow a strict rational choice framework of optimal ends-means rationality—an approach that is better suited to aggregative theories of democracy (Warren 2002). Instead, theorists of deliberative democracy generally stipulate that deliberators ought to give publicly accessible reasons for their preferences that they are able and willing to defend, drawn from a more-or-less shared world (Benhabib 1996; Chambers

1996; Dryzek 2000; Elster 1997; Gutmann and Thompson 1996, 2004; Warren 2002). Gutmann and Thompson (2004) further specify this requirement by insisting

...any premises in the argument [made within deliberation] that depend upon empirical evidence or logical inference should in principle be open to challenge by generally accepted methods of inquiry.... [and] premises for which empirical evidence or logical inferences is not appropriate should not be radically implausible (72).

Thus, the rational citizen who deliberates is not the ends-means calculative machine that some might imagine. Instead, she is a citizen who has the capacity to form, present, revise, defend, or reject preferences and judgments based on publicly justifiable reasons that are, at least in principle, subject to interrogation and evaluation by her peers, and that are not *a priori* implausible or irrelevant. These preferences and judgments should be more-or-less coherent and drawn from the world in such a way that aims to make sense of that world. As mentioned earlier, this definition of rationality is not a particularly onerous one, but it is functional and sufficient for my purposes in this dissertation.

It is important to keep in mind that for this dissertation, I am referring primarily to *deliberative rationality*, which is bound by the requirements that individuals in a deliberative process draw on genuine reasons when arguing and defend their reasons, preferences, judgments, and so on to others. This approach contrasts, for example, with ends-means rationality in which it would be rational to engage in strategic behaviour, to lie, to withhold key

information, and so forth to achieve a particular outcome. In a deliberation, this behaviour would be *irrational* since the goals of the deliberation are founded on an intersubjective and normative considerations aimed at establishing better judgments and decisions through, among other things, exchanges based on epistemically valid statements.

Let us expand on the point of deliberation prohibiting *a priori* implausible, irrational, or irrelevant reasons. A rational deliberator is concerned with *rational considerations*. In his entry on deliberation in the *Oxford Handbook of Political Philosophy*, Robert B. Talisse gives the example of Abby, who is deciding between attending a film or visiting the library (Talisse 2012: 204). She can flip a coin to decide or she can weigh the pros and cons of each option. If she chooses the coin toss, she is choosing an *internally irrational* process (which may be perfectly externally rational, if she really cannot decide). That is because the coin toss is blind to reasons; if, however, she deliberates over the decision, she can generate and offer *reasons for and against* each option: an internally rational process. Deliberative democracy is the public extension of this logic of rational deliberation; it does not presuppose specific ethical content for decisions—at least outside of very general principles related to liberal democracy more broadly. However, it does presuppose a rational and deliberate process aimed at generating coherent judgments and decisions in a formal political context. My dissertation is concerned with precisely this: how can we generate coherent, valid judgments and decisions through democratic deliberation.

## Self-determination and autonomy

Next, let us look at self-determination and autonomy, each of which I examined in greater detail in chapter two. One of the core justifications for the reason-giving requirement of deliberation is that deliberators ought to be treated as autonomous agents capable of interacting with one another towards some end of their choosing, rather than as passive objects of governance mobilized for the purposes of reaching another's end (Gutmann and Thompson 2004: 3-4). For this to be possible, citizens must be autonomous in the broader, non-cognitive sense of the term (recall that throughout the dissertation I discuss autonomy in the *cognitive sense*, not the liberty sense, unless otherwise noted)—non-cognitive autonomy refers to the capacity to freely act in such a way that allows an individual to be personally self-determining and a participant in the determination of the fate of their community.<sup>28</sup> This requires a number of stipulations in the way that deliberations are designed, carefully set up to ensure to the greatest degree possible that the autonomy of each participant is respected to the greatest degree possible; as Estlund (2008) puts it, deliberations should be set up as to prevent “power’s interference with reason” (193). Specifically, this end requires that both the deliberative design and the behaviour of deliberators be more or less in line with the goal of respecting the autonomy and agency of each participant. This end also requires institutional setups respecting the freedom and equality of each participant within the context of the

---

<sup>28</sup> For instance, this latter understanding of autonomy is grounded in an account of human intersubjectivity discussed in chapter two in the section on relational autonomy. The relational understanding of autonomy—consistent with but distinct from the one I use in this dissertation—focuses on human beings as inherently bound up and partly determined by communities, while my interest is in the *cognitive process* that either precedes or succeeds specific instances of reflection, either alone or in a group.

deliberation and the principle of non-domination so that each affected by an outcome has a meaningful chance to engage in the process that generates a decision. These requirements open up space for the possibility of self-determination—which is defined here as the possibility and practice of exercising (for now, non-cognitive) autonomy in the pursuit or realization of one’s preferences, mirroring, in a slightly confusing way, Rostbøll’s (2008) definition of *autonomy* as a process of “[living] under conditions where one can engage with others in deliberative practices that enable one continually to modify one’s preferences and opinions in light of arguments” (87).

I have been speaking of autonomy in this section in a *non-cognitive*, traditional sense. However, as noted in chapters one and two, autonomous deliberation requires the capacity for sustained autonomy in the *cognitive* sense. To review, cognitive autonomy, as opposed to autonomy as freedom, requires that “...the influences and conditions that give rise to the desire [or preference or intention] were factors that the agent approved of or did not resist, or would not have resisted had she attended to them, and that this judgment was or would have been made in a *minimally rational, non-self-deceived manner*” (Christman 1991: 22, emphasis mine). The relationship between the two senses of the term is interesting but irrelevant to the argument I make in this dissertation. For now, it is enough to note that there are at least two distinct meanings of the term “autonomy,” and to distinguish one (i.e. autonomy as freedom-directed in the non-cognitive sense) and from another (i.e. the cognitive sense). The latter adds a level of analysis to the traditional understanding of autonomy by pushing the boundaries of autonomy past the absence of constraints and into to the realm of cognitive capacity.

Before proceeding, I must address the question of whether *cognitive* autonomy, as I conceive of it here, is ever fully possible, or whether it is an ideal. As I will discuss in chapter six, full cognitive autonomy is an ideal, though I think of individual autonomy as a capacity that is exercised *in the moment* in a *better or worse way*. So, full autonomy is an ideal, but the exercise of one's autonomy falls along a continuum that can be evaluated as *more or less* autonomous to the extent that individuals are aware of what drives them. My primary concern in this dissertation is understanding how autonomy can become compromised during deliberation, since insofar as it is compromised, the validity of judgments and decisions is undermined; my secondary concern is how we can address the challenges generated by compromised autonomy and encourage better exercise of that capacity, whether or not it ever reaches the full ideal of (theoretical) total cognitive autonomy.

### **Validity of reasons**

Let us now look at the system-level assumption about the *effects* of deliberation—or the *goods* that theories of deliberation claim will be produced or enhanced through deliberation among rational, autonomous agents. Habermas (1985), summarized by Warren (2008: 184), provides a clear and precise definition of validity that summarizes well the kind of decisions that proponents of deliberative democracy contend that deliberation tends to produce. Validity, Warren summarizes, “...is an attribute of statements assessed according to whether they are factually true, normatively right and expressively sincere (or truthful).” He goes on to note that statements influence others when they are recognized as valid in the three “worlds” outlined by Habermas: the empirical (the factual world), the normative (the world of social norms and

rules), and the internal realm of feelings, desires and thoughts (the world of internal experience) (Warren 2008: 184).

The underlying force of validity rests on the quality of valid reasons and arguments—the idea that producing good decisions requires high-quality evidence and reasons (Chambers 1996), corrects or avoids mistakes while also developing understanding (Gutmann and Thompson 2004), and publicly links objects to their referents in order to generate shared understanding (Warren 2008). Validity is thus a function of the process of generating agreement about the world through the exchange of information between two or more individuals. Facts about the world are embedded in statements that are either taken up by others (or not), and thus validity is established or not in the process of coming to judgments and decisions. When it comes to democratic deliberation, *validity is not about* faithfully translating facts about the external world in the pursuit of some ontological uncovering of “reality.”<sup>29</sup> Proponents of deliberation claim that deliberative democracy often allows for the maximization of the likelihood that valid statements will be made and that a coherent shared world will be generated by those assembled to deliberate—more often than other possible decision making procedures including mere voting, technocratic directives, or coercion—since deliberation is directly concerned with *reasons* as a currency instead of other sorts influence such as impulse, expertise, or force.

---

<sup>29</sup> For more on this, see my discussion of correspondence theories of truth and consensus theories of truth in chapter two.

## More legitimate outcomes

As a process of decision making that stresses inclusion, (non-cognitive) autonomy, self-determination, reason giving, and which claims to produce more valid and reasoned outcomes than the alternative decision making processes noted above, proponents of deliberative democracy also claim that the outcomes of deliberations—or the outcomes of decision-making procedures in which deliberative democracy plays a role—are more legitimate than other approaches (Estlund 2008, Fishkin 1996, Warren 2008). However, the core justification of deliberation as a legitimate decision-making process focuses on participation and freedom. As Cohen (2002) puts it, “...free deliberation among equals is the basis of legitimacy” (91). So deliberation—whatever else it accomplishes—acts as a source of legitimate outcomes through its commitment to providing space for each to engage and to be a part of the decision-making process, so that each is bound to *a priori* to outcomes that respect the inclusive, free, and fair deliberative procedures agreed to before deliberations begin. Other goods, noted above, including more valid outcomes and rational outcomes, further support deliberation. Thus, one of the key goods generated by deliberation is a sense that one is *tied to and bound up within* the decisions generated by deliberation; the further assumption, largely borne out historically, is that the more substantive and participatory a mode of decision making is, the more it will be respected and adhered to by those who participate in it and are affected by it—which may be closely intertwined with a broader desire for self-determination.

## **From evolutionary and social psychology to deliberation: dual process cognition in the system 1 and system 2 modes**

*Consciousness reigns, but doesn't govern. – Paul Valéry*

The human brain—the site of consciousness and knowing—is not a computer. Indeed, it is *quite far from the computing machine* that it is often explicitly stated or implicitly assumed to be when we establish expectations about how individuals should engage in deliberative political discourse. In this section, I will discuss some ways in which the rational brain can help us meet the standards set by deliberative theorists and some ways in which it can, upon us finding ourselves in certain circumstances or interacting with certain institutions, work against meeting those standards. However, first, I will explore an extreme example of how a purely rational brain might operate, the brain-as-computer, to establish some parameters for when I discuss how the brain tends to operate.

Why is this necessary? Because the way we conceive of the brain is linked to how we conceive of thinking. Metaphors and analogies are powerful and it is important that we understand the fundamental architecture of thought from the brain to behaviour. Because our normative expectations are conditioned by our conceptions about how we think and how we *ought to think*, we should pay close attention to those conceptions; and since they are rooted in metaphor and analogy, it is useful to lay them bare when given a chance, to interrogate them, and, when necessary, to adjust and correct them.

The conceptual model of brain-as-computer, accepting input from the body and the environment, logically and serially computing a response, and generating a smooth output is incorrect and misleading, and I am not suggesting that serious theorists are suggesting this is how the brain functions, but it worth knowing *why it does not operate this way*. Edelman

(2006), discussed above and whom I will quote at length here, summarizes the reasons why the brain-as-computer is a poor analogy

First, the computer works by using logic and arithmetic in very short intervals regulated by a clock...the brain does not operate by logical rules. To function, a computer must receive unambiguous input signals. But signals to various sensory receptors of the brain are not so organized; the world (which is not carved beforehand into prescribed categories) is not a piece of coded tape. Second, the brain order...is enormously variable at its finest levels....no two brains are identical.

Last, it should be stressed that we are not born with enough genes to specify the synaptic complexity of higher brains like ours. [Our] gene networks...are enormously variable since their various expression patterns depend on environmental context and individual experience (21-22).

On top of the reasons offered by Edelman why the brain is not a computer, there is one more that we ought to pay attention to—one that is linked to the phenomenon of neuroplasticity. This phenomenon refers to the changes in the synaptic (and non-synaptic) connections in the brain based on the individual experiences of each. Through one's changing experiences, new synaptic connections are generated and become, along with other connections, *more or less likely* to become activated so that different patterns of neuronal firing will become more or less likely to occur based on changes in the brain brought about by varied individual experience and

environmental changes. In essence, the brain is fundamentally unlike a computer because it is *constantly changing* on an individual basis depending on the very particular lived experiences of each human being—even though shared tendencies can and do emerge, as we will see in subsequent chapters.

This means that while the brain is not a purely rational computational machine, it is not wholly subject to manipulation, either, and is malleable to the point where personal practices and institutions can direct behaviour in such a way as to assist in improving outcomes: for instance, better judgments and decisions. Conversely, as we will see below, the epistemic force of deliberative democracy as a source of good judgment and decision making is undermined if specific procedures or expectations are generated concerning the brain-as-computer model (many are not, but some are). Part of the reality of human cognition as it occurs in a complex world is that both internal processes and the external environment are complex and often enough unpredictable, and highly variable.

So, moving away from the (deliberately extreme) analogy of the brain-as-computer, what is the brain? Moreover, how does it process information for making judgments and supporting attitudinal development and change? Being able to answer this question is important, since it will guide not only our understanding of how and why individuals behave in a certain way when engaging in democratic deliberation, but also because it will provide some foundational insight into how we might design personal practices, procedures, and institutions in order to improve democratic citizenship in general and deliberation in particular.

The human brain is a bustling clearinghouse for information; it is a site for processing significant amounts of internal and external sensory data, for interpreting that data, and for

*constructing a world in our minds* from it that includes, among other things, values, and, when required, preferences, and judgments. The process of translating raw data into a coherent world and subsequent worldview, however, is not a clear and automatically high-fidelity endeavour. Indeed, the process of data management for the brain is part translation, part creation. As psychologist Daniel Kahneman (2011) notes to this effect “The world in our heads is not a precise replica of reality” (138). And when it comes to politics, as I will demonstrate in this chapter and the next, cognition, reasoning, judgment, and decision making can be particularly tricky endeavours.

But first, let us look at how specifically the brain enables the individual to make sense of the world. The leading theory about how the brain processes information for making judgments and developing or changing attitudes is grouped into a series of models of cognition known as “dual-process models.” While there are variations in these models (for a summary see Eagly and Chaiken 1993), one long-standing leader is particularly incisive: what I am calling the systems approach. This approach to dual-process models helps us understand and explain how we make certain judgments; it is most famously and cogently advanced by psychologists Daniel Kahneman and Amos Tversky.

The approach is defined by its broad division of thought into two “systems”: system 1 and system 2 modes of thinking, which I explain below (to organize four models of cognition and attitude change that I will use to evaluate the question of whether citizens can deliberate autonomously). I will sort elements of the models into system 1 (which, as we will see, is automatic, intuitive, and non-conscious cognition) and system 2 (which refers to slow, conscious, and deliberate cognition). While there are some tensions between the models and

between them and their fit in the system 1/system 2 dichotomy, they nonetheless remain consistent with the approach I am taking here. I will note these tensions as they arise. For the most part, I operate under the assumption that when it comes to democratic deliberation, participants ought to be operating in the system 2 mode as much as possible so that they are aware of the sources of their reasons (i.e. they have reasons for their reasons) and can thus operate autonomously.

The models I explore here—elaboration likelihood, system justification, automaticity, and social intuitionism—under the overarching division of system 1/system 2—cognition are essential components of the argument that follows in this chapter, and indeed in the remainder of this dissertation. From here onwards I will refer to these models individually by their particular names while referring to the systems model as a *cognitive system* or a *systems approach*. In chapter four I will focus on *motivated reasoning* in depth; I have kept motivated reasoning separate from the other models since, while it has some overlap in system 1/system 2 modes of cognition, it has some notable differences and specific implications for autonomous deliberation (since motivated reasoning is less *cognitive* and more *motivational* in most cases) that warrant separate treatment.<sup>30</sup>

---

<sup>30</sup> There are some elements of motivated reasoning that fit well with the models I will discuss in this chapter—most notably, online processing. I have included some short discussions related to motivated reasoning in this chapter, but my primary exploration of motivated reasoning will take place in chapter four.

## **The systems approach: how we generate judgments and beliefs**

What I am calling the systems approach is a two-level model of judgment (including attitude and belief formation). This approach was developed by psychologist Daniel Kahneman, alongside, among others, his academic partner Amos Tversky. It is based on decades of research in human cognition and behaviour. The system is based on two cognitive sub-systems bound up in the brain that can be grouped into system 1 and system 2. The first system—far more influential on our behaviour than often understood or admitted, and the originator of many of our judgments and beliefs—is automatic, rapid, outside of conscious control (i.e. non-conscious), and generally effortless. It is also highly subject to error and manipulation. The second system is effortful, conscious, cognitively taxing, slow, and associated with cognitive efforts associated with agency, choice, and concentration (Kahneman 2011).

It is system 2 that we tend to think of when we talk of the “self” (Kahneman 2011). However, these systems are meant to be handy metaphors—or organizational categories—rather than single variables; each system is made up of a number of brain regions and sub-processes. So, each system is made up of a stylized group of variables that serve as a helpful explanatory cognitive system (e.g. you cannot look at the brain and “see” the system 1 or system 2 regions). Each roughly accounts for a series of related psychological and cognitive processes and predispositions (Kahneman 2011: 28-30). Notably, system 1 plays an important role in generating the judgments and beliefs of system 2, often largely outside of conscious awareness. System 1 generates “impressions, intuitions, intentions, and feelings” which, if endorsed by system 2, can become adopted and integrated into more complex judgments and beliefs (Kahneman 2011: 24). And, as noted, while system 1 is generally efficient, effective, and

reliable, structural biases and shortcomings render it susceptible to errors in specific contexts, threatening the quality and integrity of both immediate judgments emerging from it and more involved system 2 judgments in which it plays a role generating and sustaining (Kahneman 2011).

Again, proponents of deliberative democracy are normatively committed, at least generally, to centering deliberations on rational, autonomous actors who can give publicly accessible reasons for their preferences. The proper functioning of a system 1/system 2 structure does not *a priori* preclude rational and autonomous deliberation. Indeed, it may *enhance* such deliberations, since the structure of system 1 is such that it can pick up on relevant information—social situational, intuitional information, including subtle cues—and use that data in the conscious reasoning around specific points. (It is important to keep in mind that system 1 still represents *a mode of cognition*.) However, to the extent that system 1 generates structural biases that interrupt or distort information processing carried out in system 2, it can undermine the integrity of the judgments that follow, threatening, when it comes to judgment, the required aforementioned capacities of rationality and autonomy.

For if deliberation is based on system 2 reasoning founded on non-conscious, structurally biased system 1 cues—whatever their origins—the autonomy of the deliberator might itself be structurally undermined. So, the integrity of deliberation can be said to rest partially on the extent to which structural, non-conscious biases find their way, via system 1 cues, into the thinking, reasoning, and, ultimately, the judgment of participants in a deliberation. It is thus important that we determine the conditions under which such biases might emerge and examine how specifically those biases might affect autonomous

deliberation.<sup>31</sup> In this dissertation I am interested in sorting cognition, broadly, into system 1 and system 2 modes for the purposes of understanding when individuals who deliberate are best able to engage in autonomous judgment and decision making; I am particularly interested in how system 1 thinking increases the chances of cognitive distortion and bias, and how it undermines an agent's capacity for autonomy and rationality as they would be expressed through particular sorts of judgment made through system 2 (which is itself subject to bias and manipulation insofar as an individual's reasoning can be undermined at any point in the thought process).

### **What is cognitive distortion and bias?**

As mentioned above, in this dissertation I will explore four models related to cognition that explain how judgment is susceptible to what I am calling cognitive distortion and bias. By "bias" I am referring to a systematic—and unreasonable—cognitive privileging of some perspective, preference, desire, etc., or of series of these that undermines rationality and autonomy.<sup>32</sup> The term "cognitive distortion" requires and deserves a bit more commentary. By cognitive distortion, I am referring to a phenomenon in which there occurs a *non-conscious transformation of the content or meaning of data during the process of cognition that would, if the individual were aware of it, be unwelcome and/or would contradict their understanding of*

---

<sup>31</sup> It is important to recall that these systems are *models for interpreting behaviour* rather than specific types of brains or brain structures or personalities.

<sup>32</sup> The term "bias" is typically meant pejoratively insofar as it is associated with unfairness; however, bias does not need to be associated with unfairness. Indeed, we can imagine bias being positive since, strictly speaking, it refers to prejudice in favour of something or someone. One can have a bias, for instance, in favour of healthy food over junk food. My concern in this dissertation is with *unreasonable and hidden* bias that undermines rationality and autonomy.

*how they make judgments and decisions.* To distort something means to change its original shape so that one is left with some product or perspective different than before the distortion. In the context of cognition, distortion occurs when some initial data or thought is transformed during the process of transmission in a way that would be unwelcome to the individual who is thinking or would contradict their understanding of how they think, judge, and decide (this is what separates *distortion* from mere *change*).

In the context of the two systems approach, the distorted or biased cognition that I am concerned with here occurs under system one processes, outside of the awareness of the individual (compared to, say, deliberately distorted cognition, as when someone *consciously* rationalizes or *deliberately* chooses information selectively—as is the case in *some instances* of motivated reasoning). It would be *unwelcome* to the extent that if an individual were exercising autonomy and processing information centrally (e.g. through system two) she would either *deny or amend* the choice or judgment she reached. Again, as I have noted, the problem at hand is not that certain information, research strategies, or cognitive practices are bound up in affective cognitive processes, but rather *that these are hidden, may be biased, and are unavailable for individual or collective evaluation during the process of establishing validity through deliberation.*<sup>33</sup>

Admittedly, there is little cognition—perhaps none—that occurs without *some* distortion or bias. As related above, the world in our heads does not seamlessly match the world outside of it—which, in part, is the why correspondence theory of truth is problematic.

---

<sup>33</sup> I am, of course, assuming individuals prefer to know their motivations and to be correct about why they reason the way they do. For this dissertation, I am bracketing the phenomena of deliberate self-delusion and lying in the contexts of democratic deliberation.

Our brains and minds are constantly actively *constructing* the world by processing external raw data and transforming it into something coherent and useful to us. However, the cognitive distortion to which I am referring in this dissertation is structural and counter-productive—at least in the context of democratic deliberation—since it undermines the quality of one’s reasoning (and undermines self-determination) by altering (outside of one’s awareness) the data, thought(s), or reasoning with which one is engaging to produce a judgment. Moreover, this cognitive distortion is *unwelcome*; similar to the definition of autonomy offered by Christman (1991), and discussed above, a key requirement of cognition that is not distorted is that the individual who is thinking would affirm any transformation of data during the process of cognition.

Whereas much of our cognitive activity is creative—again, we are constantly building a world in our minds out of the raw stimuli found in the world outside us—cognitively distorted processing, which, unlike the brain and mind’s act of assembling a world from raw data, can in principle be avoided or at least mitigated. It might be easiest to put it this way: all cognition is creative, but cognitive distortion *warps data and the process of reasoning from that data* in a way that undermines autonomy and threatens our ability to generate valid judgments and decisions. This effect is why cognitive distortion earns a pejorative status within a particular cultural context (e.g. democratic deliberation) relative to certain normative expectations (e.g. that those who deliberate are rational and autonomous).

We could, if we wanted to, agree that *any reason based on any consideration, conscious or not*, would be fine political fodder for democratic deliberation. In such a case, distorted cognition would be perfectly acceptable, since the status of our preferences and judgments

would not be subject to evaluation based on our autonomy. However, since proponents of deliberative democracy have directly or indirectly specified the need for rational and autonomous actors, cognitive distortion is obviously a threat, since it acts as a challenge to both the desired goods of rationality and autonomy. Appreciating the ways in which cognitive distortion occurs or is manifest is key to understanding the failures of rationality and autonomy to explain fractures in the democratic process. Once again, our cultural assumptions, desires, or requirements for democratic deliberation run up against certain natural phenomena and the limits of our cognitive realities in the given context of contemporary, fast-paced, complex liberal democracies. Moreover, while that challenge does not imply we can do nothing to address the problem, it does suggest that current modes of democratic deliberation may fall short of meeting certain needs and that the status quo undermines the plausibility and reduces the quality of democratic goods offered by the deliberative democratic approach.

For now, however, before addressing potential fixes to the problem of cognitive distortion, let us turn to the specific threats to autonomous, rational judgments that are explained by the four models mentioned above and see specifically how such distortion might occur and how it might affect democratic deliberation in each instance. These four models have been chosen for four reasons discussed in chapter one, but two reasons are particularly important and warrant repeating. First, they are representative of the broader literature on cognitive distortion insofar as they are generally accepted theories in social psychology and include the majority of known cognitive biases (i.e. they are representative and comprehensive). Moreover, second, they are of specific concern to democratic deliberation

insofar as they pertain specifically to capacities of rationality and autonomy, as defined in this dissertation, required for good judgment and decision making (i.e. they are case-appropriate).

## **Models of cognition**

### *The elaboration likelihood model (ELM): persuasion and opinion/attitude change*

The ELM, which is concerned with how attitudinal change is brought about in individuals, was created by psychologists John Cacioppo and Richard Petty. It has remained, more or less, similar to the form that they developed in the 1970s and 1980s. The ELM helps us understand how political judgments are formed and how they change (if they change at all). During the time in which the ELM was developed, social psychologists turned away from approaches that focused strictly on rational, deliberate opinion/attitude change and began to embrace new understandings of cognition that considered less effortful and conscious mental activity. The ELM breaks down attitude determination into two routes set as poles along a continuum measuring one's motivation for elaborated thought (i.e. thinking capable of rationally processing a complex persuasive message): a central route, which is characterized by cognitively effortful and taxing thought, and which accounts for *high elaboration likelihood*; and a peripheral route marked by the use of affect, cues, and heuristics, accounting for *low elaboration likelihood* (Petty 1999; Petty and Cacioppo 1981, 1986). We can place central processing in the system 2 category and peripheral processing in the system 1 category; and we can understand our challenge in deliberative democracy as one concerned with encouraging more use of system 2 functions to encourage central processing and high elaboration when changing attitudes and generating judgments.

For example, consider an individual who is opposed to tax increases but presented with a situation in which his local transit authority is in need of funding for infrastructure spending. Here, a specific need (i.e. transit funding for infrastructure) meets a pre-existing attitude (i.e. higher taxes are undesirable). Say also that the man regularly uses public transportation, but his experience on transit is negative because of the inconveniences he faces daily because of inadequate transit infrastructure. The man in question may have good reason to engage in central route processing around this issue since it has high salience for him. It may be worth his time to engage in effortful, cognitive taxing reflection on the issue. He may well decide that a tax increase is, in fact, necessary and, on balance, desired in this case. He would presumably have good, immediate reasons, too (e.g. commute times would improve, the bus would be less crowded, etc.). Of course, he directly benefits from this outcome, but it could just as well be the *reasons* that have changed his mind (i.e. the force of argument/reflection) rather than the fact that *he benefits from those reasons bringing about a decision*. Conversely, in the same situation, a driver who is also opposed to higher taxes, who never uses public transit, and who is not affected by long commute times, may never have good reason to examine the case for/against raising taxes to generate funds for transit infrastructure and may thus engage in peripheral route processing. To simplify, such processing might look something like this: “They want to raise taxes to spend on transit. I don’t like taxation. So, I’m opposed to increased taxation for transit funding.” Note that in the former case *reasons* are central (cued by the salience of the issue to the first man) and the in latter case reasons are absent (instead, affect is central since the man does not “like” taxes).

Within the context of the ELM, researchers find that situational and dispositional factors are most significant in determining which route an individual will take when processing a message that is intended to persuade them. Those who “enjoy thinking,” for instance, and thus have such a disposition to engage in prolonged reflection, are more likely to engage in central route processing than those who do not enjoy thinking (Cacioppo and Petty 1982).<sup>34</sup>

Situationally, real-world factors, such as time-constraints, can also effect where one falls along the continuum in any given circumstance (and, potentially, they can shape one’s dispositions in the first place), with greater constraints decreasing the likelihood that one will engage in central route processing (Moore et al. 1986). However, more importantly for my purposes here, given that central processing is cognitively costly, issues that are *relevant* to an individual and which they are *motivated* to engage with are more likely to be centrally processed than those that are not, with the latter more likely to remain processed through the peripheral route. Thus, both *personal psychological* and *broader environmental* factors are at work in determining route choice. So, when it comes to democracy in general, and deliberative democracy in particular, those with more at stake in a given decision should be more likely to engage in central processing (Heine personal communication 2015), while others might (at least initially) be more likely to engage in peripheral processing. However, as I will argue in chapter five, to improve the odds that an individual generates good judgments and that a group

---

<sup>34</sup> It does not appear to be the case that high-elaboration and low-elaboration tendencies are fixed, say, in the way that raw athletic talent is thought of as fixed. Accordingly, generating tactics and strategies for making high-elaboration more likely is an appropriate—and potentially quite productive—research objective, and one I will take up in chapter six.

produces good decisions, participants in deliberation should be specifically encouraged and motivated to engage in high-level processing.

Notably, research on the ELM also suggests that there are differential effects on attitudinal change and persistence based on where one finds their cognitive effort falling on the ELM continuum. Unsurprisingly, persuasive messages considered along points closer to high elaboration likelihood *tend to be* more durable and impactful when adopted (Petty et al. 1995) than those closer to low elaboration likelihood, though this does not necessarily account for bias or distorted thinking that may creep into higher elaboration states (a problem that will be discussed later in this chapter and again in the second half of this dissertation). The key takeaway is that the route taken when processing information can matter a great deal: the same combination of the individual, the context, and the information presented can have a different outcome depending on where one falls on the ELM continuum in any given instance. As we will see in more detail, depending on potential structural tendencies in ELM route choice, there are significant potential implications for political opinion and attitude formation, including different real-world political outcomes, for how democratic deliberations are set up and run, and for the types of judgments that emerge from them.

Obviously, proponents of deliberative democracy, especially those who advance deliberation as a means of making epistemically valid decisions, are committed to and expect (at least relatively) durable, though revisable, attitudes—again, these are among the cultural assumptions of deliberation. It is equally obvious that these proponents are likely to be just as normatively committed to encouraging more central route processing. Even though scarce cognitive and day-to-day resources make choosing to expend cognitive effort on persuasive

messages a careful art—when indeed we have some control over the choice—proponents of deliberation have good reason to argue that deliberative settings are highly appropriate locations for this expending. This is especially true given that central route processing is less susceptible to cognitive error or distortion, and deliberative forums are expressly meant to elicit honest and reasoned exchanges among participants in search of legitimate, durable decisions—that is to say that critical reflection is more likely to elicit reasons that must be defended than other, less engaged, modes of decision making. The key concern, then, becomes how we can motivate deliberators to engage in central route processing through reason giving—again, a question that will be systematically taken up and explored in the second section of this dissertation, in chapters five to seven.

### *System justification*

In social psychology, system justification is defined as the “process by which existing social arrangements are legitimized, even at the expense of personal and group interest” (Jost and Banaji 1994: 2). Evidence from decades of research into the motivation of decision making, beliefs, and judgments through the lens of system justification theory has shown that many individuals are motivated, psychologically and non-consciously, to hold favourable attitudes towards the existing social structure—the status quo—even occasionally overriding ego justification (i.e. personal interests) and group justification motives while doing so. As Jost et al. (2004) note, while there is a “general psychological tendency to justify and rationalize the status quo, we do not assume that everyone is equally motivated to engage in system justification” (912). Nonetheless, the tendency to justify the status quo despite strong reasons

to challenge it is common enough to warrant study and, by implication, concern, given that for some issues (e.g. climate change, poverty, war), the status quo can be threatening, costly, or both. Indeed, the failure to challenge the status quo is potentially catastrophic in certain instances—not to mention critical to particular populations in others (e.g. around questions of redistribution or the effects of climate change). Moreover, *advantaged groups* tend to show more implicit support for their *ingroup*, while *less advantaged groups* often show more support (which can be either conscious or non-conscious) for *outgroups*. This tendency also has an effect on judgments made by such individuals—for instance, leading individuals in outgroups to work against their group and its interests or to denigrate their group, potentially at a personal cost to their self-esteem and task performance (Batalha et al. 2007, Dasgupta 2004).

Underlying system justification tendencies, according to Jost and his colleagues, are a series of processes of motivations that generate beliefs and judgments that support the status quo.<sup>35</sup> These include ideological motives to justify the status quo that operate implicitly and non-consciously (Jost et al. 2004: 912). One of the key motivations for these tendencies is the individual's drive to hold favourable attitudes not just towards themselves (ego justification) and their ingroup (group justification), but also towards the social and political systems that surround them (Jost et al. 2004: 887). In a case study on political conservatism, Jost et al. (2003) found common traits among conservatives at personal, epistemic, existential, and ideological levels; they distilled the core of conservatism to resistance to change and the justification of inequality. Findings also indicated that while motivations to hold conservative ideological

---

<sup>35</sup> System justification may ultimately be a particular example of loss aversion/the need for ontological security/attempts and maintaining frameworks of meaning. One of the common threads in this dissertation is the pervasive human desire for stability and familiarity.

positions varied depending on particular individuals and situations, they were linked to the needs of *managing uncertainty and threat* (Jost et al. 2003). These motivations—underwritten by a process known as socially motivated cognition—reflected both social (e.g. a desire to maintain the status quo) and cognitive needs (e.g. to protect against the fear of loss and death), though these considerations do not necessarily consciously or explicitly factor into the reasoning or reason giving of individuals when they are explaining their beliefs or rationalizing their judgments. Liberals, on the other hand, tend not to engage this palliative function of system justification as much as conservatives—and thus tend to be less happy—though they do engage in system justification behaviour (just not to the same extent, on balance, as conservatives) (Napier and Jost 2008).

The presence of system justification in thinking is not, however, absolute; the effects of related elements (e.g. ingroup favoritism and outgroup favoritism)<sup>36</sup> are occasionally moderated, mitigated, or even eliminated under certain conditions in which people were motivated to control prejudice or are made aware of potential bias and given an opportunity to control it (Dasgupta 2004). Moreover, critics of system justification suggest that researchers need to show that individuals are *actively* engaged in system justification (through bias) rather than *passively* following with the status quo, though neither would be good for democratic deliberation (Rubin and Hewstone 2004; see also Kay et al. 2009). Other factors involved in dampening support for system justification include moral outrage, existential guilt, and a desire

---

<sup>36</sup> There is some debate about whether ingroup/outgroup favoritism is best explained by system justification theory or other theories. For a discussion of this see Rubin and Hewstone 2004.

to help the less fortunate.<sup>37</sup> It has also been hypothesized that feeling “a sense of power,” experiencing emotions associated with being powerful, and seeing the existing system as illegitimate, unstable, or fundamentally unfair (e.g. closed to individual advancement) may assist in bringing about resistance to system justification (Martorana et al. 2005). Still, on balance, people tend more towards rationalizing the status quo through system 1 (non-conscious) “cognitive adjustments” than they do engaging in bringing about increased equality or justice (Wakslak et al. 2007) or through engaging in other anti-system behaviours or beliefs, unless system change seems inevitable—in which case individuals engage in a rationalization of the *new, incoming* status quo (Kay et al. 2002).<sup>38</sup>

Despite some limits to the extent of the effects of system justification and notwithstanding the critiques I have mentioned, the theory is widely accepted in social (and political) psychology; thus, for proponents of deliberation, system justification theory in general, and motivated social cognition in particular should be deeply troubling for at least two kinds of reasons. First, at the level of autonomous deliberation, individuals are expected to provide reasons for their preferences and judgments. Moreover, those should be one’s *actual* reasons and should reflect *known motivations*, since if awareness of the factors that go into generating preferences and judgments—whether they be other, unknown reasons or hidden motivations—was possible, individuals might alter or reject them. Second, deliberation requires

---

<sup>37</sup> This research is specifically focused on system justification and support for redistributive social policies.

<sup>38</sup> One quickly sees a fundamental challenge to resisting the system justification phenomenon: if having power, or perceiving oneself as having power, is required to motivate engagement in anti-system activity, then it will be difficult for marginalized groups, who tend to lack power, to engage in such behaviour.

rational actors in a broad sense, but is threatened when those individuals become *rationalizing* actors. While the former concerns itself with adopting an accessible and plausible reason-giving-based approach to deliberation, *the latter undermines this by potentially obscuring true reasons and motivations and potentially leaves concerns that should be up for discussion and debate off the table*. Neither is consistent with self-determination.

The general threat to deliberation raised by system justification theory is that any deliberative procedure that is unable to account for the subterranean—system 1—motivations of individuals, or its effect on conscious reasons—system 2—risks reinforcing in some or replicating in others existing non-rational, non-conscious motivations. Moreover, it may well also tend to contribute to the maintenance of the status quo or else to become a battle over differing visions of what the status quo should be, rather than a give and take of reasons for and against outcomes or the rational transformation of preferences in the search for better-than-chance good decisions under conditions productive of epistemic validity. If social justification tendencies act as a non-conscious filter during deliberations, either at pre-consideration or post-consideration moments, then deliberative reason giving is undermined, given that we cannot assume that deliberations would proceed the same way or that outcomes would be identical if all motivations and true reasons were known. Thus, these threats to rationality, autonomy, and good judgments and decisions are serious, with potentially severe implications—especially for already vulnerable groups that are co-opted into supporting the status quo.

Even the “rational” transformation of preferences is potentially undermined by instances of non-conscious acts of system justification in this case; that is because autonomy

cannot, strictly speaking, be said to *rest on* mere reason giving and explanation. As we have seen, autonomy requires that one be in a position in which they would be willing to affirm outcomes in the face of full awareness of their motivations for generating those outcomes. It is inconsistent with theories of autonomy to accept that non-conscious motivations and reasons play a significant role in determining potentially incongruent conscious justifications. Thus, to the extent that system justification occurs within deliberations, the epistemic defense of deliberative democracy as both a fair process and a way to produce good outcomes is threatened even *before* any particular decision is reached.

Recall that epistemic proceduralism as a theory of deliberative democracy—especially in Estlund’s conception of it as a theory of good collective decision making—does not rest on a thick conception of correctness, as a full correctness theory would. Correctness theories, as Estlund notes, require that a minority should be willing to accept the outcome a majority decision given that such a decision has a high probability of being *correct*; the legitimacy of the decision is drawn from its truth status (Estlund 2008). Epistemic proceduralism, however, rests on both the moral claim that it is a *fair* procedure—all who are affected are treated impartially and would accept the procedure—and the epistemic claim that it *tends to be correct on a better-than-random basis*. This system does not presuppose particular, thick normative outcomes (e.g. what justice is or the common good), but it does presuppose that the process of democratic deliberation will be acceptable to those involved and that the outcomes generated by such deliberation will tend to be good (or as Estlund would put it with a slightly different

emphasis, “correct”).<sup>39</sup> more often than they would by random chance. For my purposes, I set aside consideration of correctness and instead substitute the pursuit of *good* judgments and decisions, defined as outcomes that are transparent, valid, and reliable.

The specific challenge leveled by system justification theory to the claim that epistemic proceduralism tends to generate good or correct decisions rests on a structural critique. If certain groups of individuals involved in deliberation are similarly and non-consciously affected by a drive to justify the status-quo, and if deliberative procedures are unable to ferret out such tendencies and to either change them or determine a way to make them more conscious, rational, and autonomous, then the decisions generated by such a procedure remain susceptible to structural distortion as system 1 considerations impact ex post facto system 2 justifications. This is because any outcomes—and the reasons that support those outcomes—are potentially anchored, non-consciously, to motivations and reasons outside of one’s awareness and consequently unavailable to the agent and her peers for rational and autonomous consideration by the group. Thus, outcomes are potentially biased internally and cognitively *both before and during* deliberative exchanges, and the whole idea of a good or correct decision is *a priori* threatened, given that its justification might rest on biased, non-autonomous motivation and argumentation. Moreover, unless democratic deliberation is designed in such a way to respond to this challenge (a subject that I address in chapters six and seven), outcomes within such deliberation may be structurally and significantly compromised.

---

<sup>39</sup> Recall that for my purposes I am equating “correct” outcomes with valid outcomes in the context of epistemic proceduralism. Again, since I am not relying on the correspondence theory of truth, my concern is not about whether an outcome “matches” the world “as it is.” Rather, my concern is that an outcome is accessible, valid, and reliable.

This concern is especially significant since some different distortions might occur throughout a deliberation, affecting different individuals in various ways at different points in time.

This threat operates on at least two levels: the immediate level and the system level. At the immediate level—or what we might call the level of accuracy—non-consciously biased argumentation undermines individual arguments in specific cases. For instance, if one argues against increasing national rates of immigration on the grounds that admitting more immigrants will undermine economic recovery, but is actually motivated by xenophobia or a fear that *any change* in the status quo is a threat regardless of what the data says, then the immediate accuracy of that argument—and any outcome it generates—is seriously called into question. Moreover, certain reasons that would need to be scrutinized to produce good deliberation based on valid judgments and decisions might remain unknown to all those concerned, and thus be off the table for review.

At the system level or the meta-level—what we might call the level of epistemic validity, concerned with what counts as correct and under what circumstances—our very efforts to establish the general parameters of accuracy might be biased in favour of the status quo and might thus preclude certain ways of thinking and arguments *before deliberation even occurs*. If that is the case, then deliberations are immediately and structurally constrained in such a way that empowers certain groups—likely those with existing power and authority, as we see with system justification explanations of behaviour—and undermines those who may already be at a structural disadvantage based on the distribution of particular biases and distortions (which may vary within and across instances of deliberation). On this count, constraining the range of what counts as correct undermines epistemic procedural claims to

tending to produce correct outcomes, or, at least, shrinks that claim's territory to a smaller subset of interests, desires, and possibilities.

For instance, research into compensatory control models has shown that support for external systems is tied to “the existential threat posed by lowered (or chronically low) levels of personal control” (Kay et al. 2008: 30). According to Kay et al., this is because external systems—for instance, religion—impose an order that is otherwise perceived as lacking in the lives of certain individuals. Moreover, when these models are challenged or threatened (implicitly or explicitly), individuals tend to grasp onto them ever more strongly (or else they find a consistent alternative), thus potentially further polarizing discourse—a phenomenon also associated with the theory of motivated reasoning that I will explore in much greater detail in chapter four. So, there is no point at which deliberation begins as a “blank slate.” Of course, few, if any, proponents of deliberation suggest that there is such a thing; but proponents of the epistemic defense of deliberation rarely discuss the prior and underlying conditions that make up the ontological reality of the participants who will direct both the conditions of deliberation and the deliberations themselves. This chapter, and specifically the model of system justification, reminds us that paying attention to such phenomena is essential to improve judgments and decisions within democratic deliberation. In chapter five I further pursue a discussion of how this might be done.

### *Social intuitionism*

The next model of cognition that I will review is social intuitionism, advanced by psychologist Jonathan Haidt. The engine that drives social intuitionism is the phenomenon of moral

intuitionism. This is the claim that moral reasoning does not *cause* moral *judgment*. As Haidt argues, moral “truths” are often *perceived interpersonally* rather than argued for rationally and in isolation (Haidt 2001). Haidt’s rethinking of moral judgment turns on rationalist model of judgment on its head. Traditionally, in the rationalist model, judgment is reached through a process that begins with an eliciting situation and which proceeds to reasoning—perhaps mediated, somewhat, by affective considerations. Finally, the process of reasoning produces a judgment (Haidt 2001: 814-815). However, in Haidt’s model of moral judgment, the eliciting situation leads to an intuition (system 1), which generates a judgment, *followed by the ex post facto reasoning (or rationalization) to support that judgment* (system 2). In the model, reasoned judgments and private reflection are hypothesized to occur only infrequently (815). What makes the model *social* is the next step: one’s reasoning and/or judgment may feed into the thinking of another, affecting their intuition, reasoning, and judgment, creating a cascade effect of intuitionist judgment among individuals (Haidt 2001: 815-816). Moral reasoning is thus, argues Haidt, fundamentally interpersonal. Indeed, such reasoning is “an ex post facto process used to influence the intuitions (and hence judgments) of other people” (814) so that the social intuitionist is “a lawyer trying to build a case rather than a judge searching for the truth” (814)—which is a clear threat to autonomy and hardly good news for proponents of democratic deliberation and its epistemic defense.

Haidt offers four counter-arguments against the claim that rational, causal reasoning motivates moral judgments: first, reasoning has traditionally been overemphasized compared to intuition—both of which are known to be cognitive processes at work in producing judgments, though in very different ways (another example of the nature/culture gap I have

discussed throughout this dissertation). Second, research in social and cognitive psychology has shown reasoning to be often motivated (for instance, see the discussion of system justification above and of motivated reasoning in chapter four). Third, our reasoning process has been shown to be as much, or more, about *rationalization* than rational consideration (see, for instance, Lodge and Taber 2013 and Sniderman et al. 1986). Fourth, and finally, to quote Haidt, “moral action co-varies with moral emotion more than with moral reasoning” (Haidt 2001: 815).

The essential point of Haidt’s argument for social intuitionism is that the process of generating moral judgments is a social process based primarily—though not exclusively and universally—on rapid, intuitive, and non-conscious evaluation (again, system 1) rather than solitary, deliberate, rational reasoning (system 2). That said, Haidt’s model allows for instances of *actual* moral reasoning. He notes that “People undeniably engage in moral reasoning. But does the evidence really show that such reasoning is the cause, rather than the consequence, of moral judgment?” (Haidt 2001: 817). As noted just above, strong evidence exists to support the argument that one’s chain of reasoning does not necessarily always proceed from reasoning to judgment, but rather can go from judgment to “reasoning.”

The rapid, affective, and non-conscious nature of this model presents a challenge to democratic deliberation, especially vis-à-vis moral issues. Haidt’s model is specifically concerned with *social* moral judgments; insofar as such judgments are social, they are also political, since social judgments and political outcomes are related; such judgments can thus plausibly be said to be the heart of deliberative democracy since it is concerned with political outcomes, especially ones concerned with morality. For while democratic deliberation can be

imagined to extend to cover almost any political issue, one of the core virtues of theories of deliberation is that their proponents tend to advocate strongly for the ability of well-designed deliberations to provide space to explore moral issues in public in a generally constructive way (Gutmann and Thompson 1996, 2004). Presumably, this line of defense includes the tendency for democratic deliberation to make space for making and sharing moral judgments<sup>40</sup> and for employing moral reasoning, defined by Haidt as “a conscious mental activity that consists of transforming given information about people in order to reach a moral judgment” (Haidt 2001: 818). All of this is fine for deliberative democracy. However, what are proponents of deliberation to make of moral *intuitions*? Haidt defines these as “the sudden appearance in consciousness of a moral judgment, including an affective valence...without any conscious awareness of having gone through steps of searching, weighing evidence, or inferring a conclusion” (Haidt 2001: 818). Again, the problem this presents for democratic deliberation is not that moral intuitions or impulses (e.g. arising from socialization or empathy) affect judgments, but rather that to the extent to which this phenomenon is present during deliberation and *outside of the awareness of deliberators*, important and relevant information

---

<sup>40</sup> Most political issues—perhaps all of them—include moral judgments in one way or another. In this case, I am distinguishing moral judgments (judgments about what is “good” or “bad” relative to a given value system) from non-moral judgments about what is most efficient, likely to work, be acceptable to a group, and so forth. For instance, it is a moral judgment to conclude that legalizing marijuana is appropriate (say, because we want to maximize liberty) and a non-moral judgment to conclude that such a decision would increase state tax revenue decrease the prison population.

required for developing and justifying reasons, judgments, and decisions may be left off the table.<sup>41</sup>

There are critiques of the social intuitionist model. One important critique of Haidt's work comes from an alternative social intuitionist model in which explicit moral reasoning—instead of intuitive leaps—is more common; indeed, such reasoning is taken to be part of “moral common sense.” In this alternative model, the role of one's moral *reasoning* in directly affecting another's thinking is considered (Paxton and Greene 2010). This critique is important because it opens up space for collective moral *reasoning* rather than intuitive leaps supported after the fact by *rationalizations*. This model, however, is more of a modification of Haidt's own than an entirely new model itself, and in going forward, especially in chapters five to seven, where I suggest approaches to overcoming challenges posed by social intuitionism and other problematic cognitive phenomena, I will keep this critique in mind.

Saltzstein and Kasachkoff (2004) are less sympathetic in their critique of Haidt's model. They claim that Haidt overstates the extent to which individuals rely on intuitive over deliberative “thinking” when reasoning on moral matters. They also claim that reasoning itself might play an adaptive evolutionary role—and so may itself be a part of our evolutionary heritage when it comes to cognition surrounding moral matters—and so Haidt, therefore, overstates the role of automaticity in moral reasoning as it relates to our cognitive evolution. They also argue that Haidt's model does not account for how or why our intuitions *change* under conditions of reasoned argument. Moreover, finally, they claim that certain cases studied

---

<sup>41</sup> Even Kant saw moral reasoning as a process of *testing motivations*. My concern here is not with the origins of moral sentiment, but rather one's ability to access and share their motivations during deliberation.

by Haidt were not cases of intuitive leap, but, in fact, reasoning. These critiques are important, but given the bulk of evidence that supports the social intuitionist model, they serve more as checks *on the extent to which the model is present and active in everyday life* rather than as either an alternative account of human cognition in social contexts or sufficient reason to abandon the social intuitionist model all together. It is not surprising that Haidt's model is not the only story about how we generate moral judgments—no model tells the whole story. However, Haidt's model sheds light, at the very least, on how some of us (likely many of us) make moral judgments and must be taken seriously.

So, despite these critiques of Haidt's model, the challenge presented to deliberative democracy by cognitive phenomena explained by the social intuitionist model is a potentially significant one. This challenge is that autonomous deliberation is undermined to the extent that rapid, non-conscious, intuitive moral judgments are generated by an individual and then supported by ex post facto reasoning—or rationalization. To be clear: *it is not merely the fact that these judgments are intuitive that is the problem*. Intuition can be a valuable mode of cognition and source of judgment; indeed, intuition can access certain evaluative criteria that may be outside the reach of rational considerations or evaluation. The problem is that such intuitions usually occur *outside of awareness*, and so they are difficult to evaluate collectively in a deliberative setting (since the individual may not have access to them). Moreover—and perhaps more problematically—this mode of “reasoning” subsequently tends to lead to the individual searching for ex post facto rationalizations for a judgment made based on considerations and motivations outside of their awareness; in so doing, it undermines autonomy and poses a *structural* (i.e. regularly embedded into the process of deliberation at

each stage) threat to deliberative democracy as a epistemic theory of making public political judgments for much the same reasons noted regarding system justification.

Recall that the epistemic dimension of deliberation requires autonomy to generate valid outcomes. As Warren notes, validity is established *procedurally* and *interpersonally*—achieving validity requires that others recognize a statement or judgment to be accurate, true or correct. In politics, it is not enough to merely appeal to some personal ontological status as the source of validity (Warren 2002). However, Warren adds that in the case of “...needs, desires, preferences, and experiences...the validity of assertions depends upon convincing others that...one is representing oneself as truthfully and sincerely as possible” (Warren 2002: 193). In either case, however, some form of public reasoning is required. For deliberation to proceed as an exercise in making judgments and reaching decisions based on accessible reason giving, it is not enough for X to assert that they prefer Y. In the deliberative context, X must give a coherent reason or series of reasons that can plausibly be expected to be evaluated as genuine. These requirements imply that deliberation is a process of *public reasoning*, not just *rationalizing* pre-determined, non-conscious, and inaccessible preferences as judgments. The ability of individuals to meet these requirements is potentially undermined in cases in which intuitionism is driving judgments in the system 1 mode outside of the awareness of the individual who is participating in democratic deliberation.

But what difference does it make *where* one’s reasoning comes from if it *appears to be* autonomous? There are at least three reasons why this is problematic in a deliberative setting—and likely elsewhere, too. First, if autonomy is undermined by rapid, non-conscious, and affective judgments at the system 1 level, then we must admit that our control over our

judgments is seriously limited in some cases, if not eliminated altogether. In any cases in which the work we do on our moral judgments is that of rationalization rather than causal reasoning, then we cannot be said to be meaningfully in control of our judgments, at least *in the moment*.<sup>42</sup> A loss of autonomy and a reasonable level of causal control limits the range of possible outcomes available at the outset, thus potentially limiting the range of options that are placed on the register of legitimate possibilities.

This leads us to the second, related, challenge revealed by intuitionism. If judgments are caused by rapid, non-conscious intuitions, and if we tend to sometimes rationalize rather than rationally-generate and review subsequent judgments, then it is possible that certain outcomes would be different if we had more autonomous control over our moral judgments in the first place. Deliberative democracy is about reason giving and preference formation/transformation. The vehicle for each of these endeavours is, primarily, talk. However, to the extent talk is undermined by intuitive moral judgments, then it is rendered less effective or perhaps even useless in some cases. By undercutting the power of talk and the force of argument, rapid, non-conscious, and intuitive moral judgments may minimize or eliminate ranges of legitimate and potentially strong judgments and decision options while reinforcing the preservation of the status quo. This might lead to deliberation being ineffective, inconsequential, or even being used as mere democracy washing.

Third and finally, and related to the previous point, rapid, non-conscious intuitive moral judgments that are supported by rationalizations undermine attempts to establish validity and

---

<sup>42</sup> If long-term affective and rational work can be undertaken to change the tendencies of our rapid, non-conscious moral judgments, then the case may be different under those circumstances.

authority through interpersonal, deliberative engagement. The deliberative process is about *honest* exchanges, while intuitive moral judgments, to return to Haidt, are closer to “a lawyer defending a client than a judge or scientist seeking the truth” (Haidt 2001: 820). This analogy is important. The word *defense* is particularly important. These rapid, non-conscious moral judgments are made and then *defended*, rather than reached through reasoning and then *supported* with (accurate) reasons. The individual is in defense mode, and thus less likely to be open to transforming their preferences. Again, our cultural assumption of autonomy runs up against our natural inclination to be sometimes driven, outside of our awareness, by biological imperatives and processes. As Haidt notes, the motivations that drive our moral intuitive judgments have likely been deeply encoded through evolution and concerned with, first, *relatedness motives* to manage our impressions and maximize smooth social interactions; and, second, *coherence motives*, which are concerned with our ability to sort and make sense of a complicated world and our particular (cultural) understanding of it, however we are able to do so (Haidt 2001: 820-822).<sup>43</sup>

The social intuitionist model is not, however, an anti-rationalist model; rather, it is a bridging model between rational and a-rational modes of cognition. As Haidt argues, the model takes into consideration “reasoning, emotion, intuition, and social influence” (Haidt 2001: 828) in an attempt to reorient cultural assumptions and expectations, drawing them closer to a balanced view of cognition and reasoning. This attempt is important to theorists of

---

<sup>43</sup> See also Anthony Giddens (1984: 75-78) on “ontological security” and Heine (2006) on “meaning maintenance” for other discussions about the psychological importance of making sense of the world and maintenance a stable and coherent worldview. I will address both of these concepts in chapter six.

deliberation, since given the assumptions and expectations made by many, perhaps most, theories of democratic deliberation, and especially by those concerned with epistemic outcomes, theorists of deliberation will be faced with an important decision over how to respond to the apparent gap between expected outcomes and human capacities. Meanwhile, at the same time, the potential threat to epistemic theories of deliberation remains in the form of eroded autonomy.

An erosion or lack of autonomy is not the only issue that emerges from the threat posed by intuitionism; the phenomenon of intuitionism also compromises the ability of individuals to produce judgments upon which others can reasonably be expected to rely. Such expectations are essential to deliberative democracy in particular but also to democracy in general; indeed, this reliance points to the presence of intersubjectively established markers that are used to generate trust, to coordinate action, to stabilize preferences, to motivate future exchanges, and to secure compliance. To the extent that these markers are compromised or disappear, the foundation of democratic deliberation—and perhaps democracy in general—will be compromised.<sup>44</sup>

### *Automaticity*

The final model of cognition that I will review is grounded in an account of the phenomenon of automaticity. Automaticity, a model developed by psychologists John Bargh and Tanya

---

<sup>44</sup> See Brandom 1994 for a discussion of “information chains” and “deontic scorekeeping,” concepts that he uses to address why reasons are important to democratic societies. In short, reasons enable the sort of democratic goods noted above (e.g. justification, future adjustment, justification).

Chartrand, refers to automatic, non-conscious self-regulation and motivation; and as a phenomenon it affects decision making, moral judgments, motivations, and goal pursuits, just to name a few of its most salient implications (for this dissertation) (Bargh and Chartrand 1999, Bargh 2007, Bargh et al. 2012). Automaticity, which has in the last thirty years “permeated nearly all psychological domains” (Bargh et al. 2012), is divided into three streams: perception and its effect on action, the activation of goals and their pursuit, and experiential self-evaluation (Bargh and Chartrand 1999). Much of what drives automaticity is the phenomenon of priming, which can occur outside of awareness and occurs when some stimulus activates meaning networks and cues knowledge of social structures in such a way that impacts subsequent considerations or judgments (Bargh 2007, Dijksterhuis et al. 2007).

As Bargh and Chartrand note, the existence of automatic, non-conscious cognitive causation in human behaviour is not debated; rather “...the mainstream of psychology accepts both the fact of conscious or willed causation of mental and behavioural processes and the fact of automatic or environmentally triggered processes. *The debate has shifted from the existence (or not) of these different causal forces to the circumstances under which one versus the other controls the mind*” (Bargh and Chartrand 1999: 463, emphasis mine; see also Bargh et al. 2012). Thus, the key question regarding automaticity, especially as it relates to autonomy and deliberative democracy, is: How, if at all, do automatic, non-conscious processes of behaviour and self-regulation affect individuals while they deliberate? (Or before they deliberate, if there is some pre-deliberative cognition that directly impacts deliberation.) This question is of particular importance to the phenomena of goal activation and pursuit, as we will see, and it

will be explored below. However, first, more needs to be said about what precisely automaticity is.

As I mentioned earlier, automaticity is broken into streams. The first is the action-perception stream. This stream refers to the claim that there is a link between one's perception of an action and carrying out that action oneself. As Bargh and Chartrand argue, while one's thinking is generally under one's conscious direction, perceptual activity is not, and so environmental information processing—over which one has very limited control, if any—leads to mental activity outside of the control of the individual and “introduces the idea of action” (Bargh and Chartrand: 1999). This allows for a two-stage process by which the external environment may non-consciously direct behaviour. The first step is automatic perceptual activity based on the environment; the second is the activation of behavioural tendencies through what's called the “behaviour-perception” link—the link between thought and spontaneous action based on immediate perception and encoded responses (e.g. stereotyping triggered by, for instance, gender or race). The link between one's environment and the non-conscious generation of automated behaviour can be summarized as proceeding in the following order: environment—perception—cognitive activity—behaviour (Bargh and Chartrand, 1999: 468).

The second stream of automaticity is the automatic and non-conscious generation and activation of goals and motivations. Since the mid 2000s, this area of research into automaticity has grown significantly and findings have reinforced the claim that our goals and motivations themselves can be activated outside of our awareness, and that, subsequently, after we have begun to pursue these goals or act on these motivations, we remain unaware of their origins

(Bargh et al. 2012). Typically, we think of our goals and motivations as the conscious products of reasoning and choice. For instance, I am motivated to be healthy; I have chosen that as a value and a desired end, so I choose to get up early and go for a run. I choose salad over French fries (in theory). However, what if it is our environment that activates some of our goals? Bargh and Chartrand argue that this can occur through the consistent pairing of external events and internal processes—much like the old neuroscientific adage that ‘neurons that fire together, wire together’. In such a case, conscious choice “drops out” as Bargh and Chartrand put it (1999: 468), given that reflection is no longer required and a non-conscious, automatic processes is more economical (i.e. there is little or no system 2 review, a process that is cognitively taxing). This dropping out can be done intentionally, through conscious skill acquisition—e.g. learning to play the guitar—or unintentionally, through repeated choices in specific situations. This process is known as the automatization.

Through frequent and consistent pairing, non-conscious patterns can develop whether one *intends to* or not. For instance, think of an intentional choice to learn how to play the guitar or to ride a bike. Now, compare those choices to a (hopefully) non-intentional tendency to throw recyclable material into the trash bin or perhaps onto the ground. As Bargh and Chartrand put it regarding how unintentionally instantiated instances of automaticity occur and lead to goals and motivations being activated outside our awareness: “These processes also become automated, but because we did not start out intending to make them that way, we are not aware that they have been and so, when that process operates automatically in that situation, we aren’t aware of it” (Bargh and Chartrand 1999: 469). To summarize: the link between motivations and behaviour conducive to meeting related goals can become

automatically encoded and forgotten, thus potentially, at some point in the future, bypassing one's conscious will and conscious choice (Bargh and Chartrand 1999: 469).

The third and final stream of automaticity is automatically-generated subjective experience through emotions, moods, evaluations, and judgments. As Bargh and Chartrand (1999) note, there is little controversy in asserting that emotions are generated non-consciously and that moods also generally fit into this same category, since very few of us can say that we always consciously choose our moods (otherwise why would so many people find themselves in bad moods?) (473). However, as they argue, judgments and evaluations are usually thought of as being consciously and rationally determined, especially in reference to valence evaluations (e.g. whether something is good or bad). As the authors put it "a substantial body of evidence has now accumulated that one's evaluations often (if not usually) become activated directly, without one needing to think about them, or even be aware that one has just classified the person or event as good or bad" (Bargh and Chartrand, 1999: 474). This phenomenon is known as the perception-evaluation link, and it operates on a system 1 level.

This link works as a kind of tagging system in which one automatically and non-consciously evaluates individuals or events, storing those evaluations, and then retrieving them automatically as needed when future, similar situations arise. This system is reminiscent of both Damasio's somatic marker hypothesis (1994) in which emotional cues based on past analog experiences influence behaviour, and the "on-line" model of information processing often discussed in political psychology. In the on-line model (Lodge, McGraw, and Stohl 1989; Lodge and Taber 2000), individuals store their political judgments or evaluations as affective markers and then retrieve overall valence impressions in the future, when required. In this model,

individuals keep an “affective tally” of issues, events, political groups or parties, and individuals, and update this tally as necessary by turning newly acquired factual information about them into affective tags, and then jettisoning that factual data when it is no longer required to make evaluations in order to produce judgments. Other recent research into automatic processes include facial perception and social judgments—the immediate, automatic, and lasting (e.g. anchoring) appraisal of an individual based on short-period exposure to their faces (Bargh et al. 2012; see also Ballew and Todorov 2007).

What unites the three models is the automatic and affective registration of an evaluation (its “tagging”). Bargh and Chartrand (1999) press the implications of this cognitive process, evaluating its “downstream” effects. One potential area of impact is on moods. Their experiments on subliminal priming determined that “mood was found to be a direct, increasing function of the evaluative nature of the subliminally presented stimuli...as predicted, whether a person is [non-consciously] making mainly positive or mainly negative evaluations within the current environment plays out in changes in his or her mood” (475).

But how, if at all, do these automatic evaluations affect *conscious judgments*? Or, as Bargh and Chartrand (1999) put it, “How do immediate, automatic evaluations impact on deliberate, conscious judgments about the same person, object, or event? The former will occur temporally prior to the latter. Does the automatic influence the conscious?” (475). That last question is the essential one, and, in a sense, is at the heart of this dissertation. If democracy in general and deliberative democracy in particular are about transmitting private data—personal preferences, values, vote intentions, and so forth—to others in order to generate public outcomes, and if the content of private data is determined automatically and

unconsciously, then, at least to some degree, our expectations about the nature and implications of the development and the transmission of that data need to be adjusted.

As noted above, immediate perception and social judgment have been subjects of research within automaticity. Bargh and Chartrand cite Ambady and Rosenthal (1992), who found that subjects asked to evaluate an individual on metrics including effectiveness at their job, quality of job practice, vote prediction, and assigned an evaluation period between 3 and 300 seconds, tended to produce the same predictions, independent of evaluation time. The conclusion drawn by the researchers is that at least some of our judgments of individuals are made instantly and tend to hold over time.<sup>45</sup> In some cases, in fact, longer exposure time produces *less accurate and predictive* judgments (Wilson and Schooler 1991). Closer to politics-proper, these conclusions are similar to those of Ballew and Todorov (2007), who found that citizens can make rapid judgments of competence when they are asked to view the faces of gubernatorial candidates. In fact, observers can make accurate predictions about the winner of a race—by judging him to be more competent—after 100 milliseconds of exposure to the faces of the two major candidates and with a 2-second judgment window. (More exposure time did nothing to improve the predictions, and deliberation *reduced* predictive accuracy.)

It is clear, then, that automatic, unconscious judgments affect our conscious deliberations—and at times or override them. It is also clear that, at least in some cases, deliberation (in a general sense) might work against our judgment.<sup>46</sup> As Bargh and Chartrand

---

<sup>45</sup> See also Albright, Kenny, and Malloy 1988.

<sup>46</sup> I am referring to deliberation in general and not deliberation as it relates to properly-constituted deliberative democratic procedures. The question of the efficiency and effectiveness of the latter will be revisited in the following chapter.

(1999) suggest, rapid, non-conscious judgments are a “continuous activity” with “real and functional consequences, creation behavioral readinesses within fractions of seconds to approach positive and avoid negative objects...serving as a signaling system for the overall safety versus danger of one’s current environment” (475-476). Evolutionarily, this phenomenon is more of a useful capacity than a liability, given that for much of the existence of the human species, rapid, automatic judgments were lifesavers—fast reaction time was essential and part of daily life for hunter-gatherers. As philosopher Joseph Heath puts it, evolution is “a tinkerer, not a designer” and it operates on a conservative, “good enough” basis (Heath 2014). The long-term evolution of certain capacities—such as the capacity to make rapid judgments about one’s environment—helped ensure species survival; physiologically, successful evolutionary outcomes that were located in the brain became piled on top or folded around one another rather than overhauled. Moreover, more recent, rational, language-based mechanisms are made to co-exist with more ancient, intuitive, and non-rational ones (Heath 2014: 42-48). What we end up with is a system that not only *is not “designed”* for the functions that it is asked to carry out but one that, in many ways, is poorly and inefficiently suited for those tasks (e.g. rational deliberation). However, as Heath points out, human beings, while not particularly adept at raw computing, are remarkably good at developing and using tools to control their environment and adapt it for their purposes (Heath 2014: 60-83). Thus, the problems for autonomous, rational deliberation revealed by social psychology may not be intractable.

## **Now what? Options for responding to challenges to autonomous deliberation**

Given the serious challenges to autonomous deliberation presented in this chapter, proponents of deliberative democracy as an approach to generating, among other goods, epistemically good decisions, based on individual rational and autonomous (i.e. good) judgments, are neglecting key psychological insights in their maintenance of the status quo. There are at least four ways to respond to the critique I have offered. The first option is to narrow the scope of what is expected, both regarding individual autonomy and epistemic outcomes—constraining the standard of good outcomes based on the limits of current cognitive tendencies. However, given the significance of the decisions that must be made in the coming years and decades, especially significant collective action problems such as climate change and declining voter turnout, this seems unlikely to improve democratically generated outcomes. The second option is to do nothing, to leave things as they are, which will likely result in the maintenance of the status quo, including pervasive inequality and other structural challenges (such as those just mentioned). However, this option is problematic for the same reason as the first. The third option is to drop the epistemic argument for deliberative democracy, relying instead on its ethical and political functions. This would be a mistake since a significant—though not unique—defining feature of democratic theories of deliberation is the tendency for deliberations *to generate good outcomes* (i.e. valid outcomes that serve as both immediate resolutions or steps towards a resolution, and which also serve as markers for future action). The fourth option—the one I think we should pursue—is to focus on individual capacity development and careful institutional design to maximize the probability that epistemically good decisions are made through democratic deliberation. As I argued above, while it is unlikely there will ever be a

guarantee that *all* decisions reflect the best possible outcome, it seems possible that the gap between our capacities and our expectations can be closed through improving outcomes through personal practices, deliberative design, and by building a deliberative system.

## **Conclusion and Summary**

As I have argued in this chapter, the challenge of cognitive distortion as exemplified by the four models I have discussed, and as it relates to deliberation, is a particularly strong example of the fundamental divergence between nature and culture that I have been focusing on: between, on the one hand, the evolutionary reality of our cognitive structures and the capacities they enable—our “nature”—and, on the other hand, our social, political, cultural, and technological systems, including the normative expectations they produce—our “culture.” After all, the human brain has changed very little in the last 200,000 years or so; however, our culture has changed a great deal. Again, as Kuklinski and Quirk (2000) argue, in presenting their skepticism towards the idea of a rational, informed citizenry: cognitive biases and distortions have been hard-wired into the human brain through a long history of evolution; and while, as the authors suggest, such distortions are likely to have served a function at some point in our history (and perhaps still do in some ways), they present challenges to citizenship. Recall their conclusion, quoted above that “researchers should not presume that any feature of human cognition is well adapted to the tasks of citizen” (Kuklinski and Quirk 2000: 165). Again, we have evolved for living in a prior evolutionary period; our imaginations, our tools, and our systems have in some ways outpaced our own capacity to live within the environment we have created through them. Therefore “...we cannot assume that the cognitive processes people use in making a particular

political judgment are well adapted to that use. To the contrary, if close observers see such processes as irrational or misleading, they probably are” (Kuklinski and Quirk 2000: 166).

So, the outpacing of nature by culture has led to the challenge of generating citizens capable of tasks such as deliberation (for others who make this point in different contexts, see Kahneman, Slovic, and Tversky 1982; LeDoux 1996; Pinker 2009). To put it another way regarding the inherent biological capacities of humans: imagine owning a computer for 200,000 years, during which time you are constantly updating the software, but never the hardware. While the brain is not a computer, and does not function as one, this analogy highlights clearly the problem of the brain in contemporary democracies: you have a system that adapts very slowly, but which is asked to do things that it has not specifically developed for and is not particularly well suited to do (Heath 2014). Such is the nature of the challenge of autonomous deliberative citizenship and good, correct epistemic outcomes considering the challenges explored above.

However, it is worth considering that while certain human capacities are inadequate or undermined when it comes to certain normative expectations about how we should behave, we cannot always assume maladaptation. Granted, some *biological capacities* might be limiting insofar as our evolution has not (yet?) *selected adaptations for the role of citizen*, but part of what makes human beings exceptional is our capacity to extend our rationality into the world through *external adaptations* (see, for instance, Heath 2014). By adjusting the world outside of us, we can survive and even thrive in an environment and in circumstances that may not be hospitable or particularly conducive to our success. In fact, this point is central to a secondary argument I make this dissertation: that we can develop ways to make better judgments and

decisions; if adaptation were not possible, then it would make no sense to argue that we could (soon) do better than we currently do.

I have also argued that the capacity of autonomy required for deliberative democracy presented in chapter one is potentially undermined in certain instances by certain a-rational cognitive processes. Specifically, as noted above, I use four leading models of human cognition that fall under the system 1/system 2 dichotomy: Petty and Cacioppo's Elaboration Likelihood Model, Jost's system justification, Haidt's moral intuitionism, and Bargh and Chartrand's automaticity. I have explained why I have chosen these models and how the threats to autonomous judgment they raise undermine conceptions of deliberative democracy as a moral theory founded on rational reason giving, the force of the better argument, and the possibility of translating true individual preferences into political judgments in the production of epistemically good and correct outcomes. However, this does not mean that the possibility of autonomous democratic deliberation is impossible or even improbable. Indeed, as we will see in chapters 5-7, there are personal practice and institutional design solutions that may serve as significant fixes and serve as a bridge between our natural capacities and cultural expectations. In the following chapter, I turn to the task mentioned above of reviewing these models and linking them to the broad phenomenon of motivated reasoning, especially as it relates to findings in political psychology that challenge the idea of a rational, autonomous deliberative citizen.

## Chapter 4: Motivated reasoning and democratic deliberation

The four models of cognition outlined in chapter three and examined under the two-systems-approach provide reason enough to question the extent to which theories of deliberative democracy, on their own terms,<sup>47</sup> can account for the gap between the expectation of rational, autonomous deliberators who can generate epistemically sound judgments on a better-than-chance basis and growing evidence that human cognition is often incompatible with those ends. Moreover, while each model presents challenges to the deliberative ideal of rational, autonomous deliberation, one cognitive mode in particular presents a significant challenge to autonomy, and thus deserves a thorough examination. In this chapter, I will argue that motivated reasoning presents a significant, specific challenge to theories of epistemic deliberative democracy that claim deliberation produces good and correct outcomes. Drawing when needed on the four models presented in chapter three, I will discuss motivated reasoning and its impacts on democratic deliberation. I will examine the core elements of motivated reasoning and will also discuss related cognitive and behavioural phenomena within the context of democratic deliberation and other political contexts.

Motivated reasoning is a burgeoning field of study and debate over the origins and effects of this mode of cognition are far from settled. However, there are reliable, generally accepted findings. In this chapter, I will explore some of the commonly accepted elements and effects of motivated reasoning. Ultimately my interest in this chapter is in drawing attention to

---

<sup>47</sup> My dissertation is primarily concerned with the *internal logic and justifications of* democratic deliberation. That is to say that while deliberation, on the face of it, is clearly preferable to some alternative approaches to decision making (e.g. coercion) and often preferable to others (e.g. mere aggregation), my focus is on claims about deliberation.

*the problem of generating good judgments as it relates to motivated cognition* and not to fully adjudicate between competing conceptions of motivated reasoning and the various sorts of supporting evidence for each. Where there is controversy, I will note it; I will also address the relevant critiques of theories of motivated reasoning and discuss various approaches to interpreting the phenomenon. That said, one of the models I will rely on heavily—but not exclusively—is Lodge and Taber’s (2013) model of the rationalizing individual to outline the elements and implications of motivated reasoning. I have chosen to focus a bit more on Lodge and Taber’s model because it is the most complete, recent, and *explicitly political* model of motivated reasoning currently available in the literature. The model is also well supported by years of research in political and social psychology, though it is still incomplete, faces its own challenges, and is subject to critiques that I will note throughout this chapter. I will also rely on past, competing, or complementary approaches to motivated reasoning that are not expressly political. Moreover, I will discuss elements of motivated reasoning that cut across the model, overlapping with the models I discussed in chapter three. These elements have specific implications when understood in the context of motivated reasoning, which is why I have addressed them both in chapter three and here.

In this chapter I spend a considerable amount of time exploring the phenomenon of “hot cognition” (i.e. in essence, cognition driven by emotion) *as it relates to motivated reasoning* (it is a phenomenon distinct from motivated reasoning) to demonstrate how a reversal of the “reasoning chains” of deliberation—proceeding from judgment to deliberation, rather than from deliberation to judgment—poses a threat to deliberative democracy. This is a particularly important threat since it cuts to the heart of what deliberation is about—providing

reasons and arguments for one's preferences. This critique, however, is not meant to be a justification for the wholesale disavowal of deliberative democracy as an epistemic theory of judgment and decision making. Instead, it is meant as a call to generate, explore, test, and put into day-to-day use ways to bring out deliberative practices and institutional setups that might serve to counteract, where necessary, appropriate, and possible, the potentially deleterious effects of motivated reasoning on deliberation (ideally within a deliberate system, as I discuss at length in chapter seven). After all, only by taking individuals *as they are but with an eye to what they might be* can theorists work to generate appropriate approaches to maximizing good judgments and better collective outcomes. As we will see in this chapter—and section two of this dissertation, in which I develop or build upon several approaches to overcoming the challenges explored here—closing the nature/culture gap laid out in chapters one and two requires that motivated reasoning be taken seriously and respected as an inherent, permanent cognitive phenomenon to be leveraged for better political outcomes when possible and otherwise mitigated, rather than a minor theoretical inconvenience to be assumed and then argued around.<sup>48</sup>

## **Motivated reasoning**

### *Overview*

Motivated reasoning is a cognitive phenomenon in which one's judgments are systematically biased by goal-oriented (i.e. directional) considerations that are affectively driven by beliefs and

---

<sup>48</sup> This is especially true given that political conflict probably selects for motivated reasoning.

feelings that are easy for the reasoner to access. This access occurs through a process that often, though not always, takes place outside of one's awareness (Druckman 2012). Dunning characterizes conclusions based on motivated reasoning as "shaped not only by the evidence the world provides to them but also by motivations, goals, needs, and desires internal to the reasoner" (Dunning 2015: 778). Individuals, Dunning notes, reason because they have some conclusion they want or need to reach; "Often," he argues "an important concern energizes them to deliberate toward a right answer" (778). However, the "right answer" is not necessarily the "correct" or valid or best answer to the question or challenge at hand; rather, it is the answer the reasoner *wants to* or *needs to* reach. This can occur, as I will note further below, either consciously (through system 2) or unconsciously (through system 1).

Motivated reasoning leads individuals to pursue particular, self-serving conclusions and often to rationalize those conclusions after the fact; and so, those who engage in motivated reasoning generate judgments that are closer to a *defense* of some conclusion in the style of a lawyer arguing a case than a judge applying the law (recall Haidt and social intuitionist model) or, even better, a scientist searching for the facts to either support or disprove a hypothesis. Dunning separates the motives for motivated reasoning into three sorts: epistemic, which refers to the need for individuals to establish and maintain a coherent worldview (see also my discussion of meaning maintenance and ontological security in chapter six); affirmational, which relates to the need for individuals to believe that they have control over their affairs and their world; and social-relational, which is linked to an individual's need to think of the world as fair and organized coherently (Dunning 2015: 778).

Within the literature on motivated reasoning, there are two modes of thought. One is an on-line approach to information processing in which individuals process and update information about issues or candidates on the basis of a “running tally” of affective markers. In reference to political attitudes—but holding for political information processing more generally—Bizer et al. (2006) define on-line processing as occurring when individuals “...evaluate individual pieces of information as they are received and integrate these evaluations into an overall attitude by the time processing terminates...when a judgment is required, an individual simply retrieves the overall evaluation that has been already formed” (646). This tally is kept, in part, outside of the awareness of the individual, and one of the key tendencies of this cognitive mode is to tend to generate the confirmation of existing preferences, which are usually given cognitive priority, and discounting or ignoring disconfirming information (in cognitive science this is known as “confirmation bias”) (Lodge and Taber 2013, Redlawsk 2002). Motivated reasoning in the on-line mode thus entails “systematic biasing of judgments in favor of one’s immediately accessible beliefs and feelings...” (Lodge and Taber 2013: 35-36). When motivated reasoning occurs, the “*motivation*” is to *confirm existing beliefs*.

Motivated reasoning may also include a second mode of thought—memory-based reasoning. According to Bizer et al. (2006), memory-based reasoning “...involve relatively less on-line evaluation...when attitudes are formed in a memory-based fashion, information is not evaluated as much as it is received; rather, it is stored in memory. When a judgment is required, individuals retrieve as much information from memory as they can, evaluate the individual pieces of information, and then synthesize these ‘mini-assessments’ into a global

evaluation based on that retrieved information” (674). This mode of reasoning is consistent with system 2 cognition, though it is still subject to bias.

The broader literature on motivated reasoning separates the *goals* of reasoning from the *process* of reasoning (see Hill et al. 2013 and Kim and Garrett 2012). Kim and Garrett (2012), in fact, argue that individuals engage in *both* online and memory based processing *simultaneously*, suggesting that individuals are “flexible processors” (see also Uleman et al. 1996). However, regardless of whether individuals engage in reasoning in one or both modes of reasoning (simultaneously), it is well established that on-line (system 1) motivated reasoning is common. This on-line model of information processing is typically contrasted with the Bayesian model of information processing in which individuals “coolly [consider] new information in light of prior preferences and accurately [update] those preferences by lowering evaluations upon encountering negative information and increasing evaluations when learning positive information” (Redlawsk 2002: 1021-1022).

The on-line model represents a partly rational process, insofar as information is being processed in accordance with a plausibly rational rule (e.g. to prefer people who are like you). However, thinking within the boundaries of the on-line model is also highly subject to bias and distortion, and, as we have seen, such a mode of cognition tends to obscure data that could be essential in developing a rational judgment that runs contrary to what one’s intuition might otherwise suggest, taking it off the table or making it harder to access in future considerations. For, while emotionally-driven, or emotionally-tinged, cognition is important for processing information and generating judgments, as I will argue in depth below, when it is distorted (e.g. in cases of deliberate or accidental manipulation), it may lead to judgments one would not

make or affirm if they are relying primarily on the system 1 mode of thinking. The on-line phenomenon undermines autonomy and the epistemic defense of democratic deliberation.

### *Hot cognition*

Motivated reasoning, as noted, is driven in part by affective considerations that are typically, though not universally, processed outside of the awareness of the individual. As mentioned, hot cognition is, essentially, an affective mode of cognition in the system 1 thought mode. Hot cognition produces judgments that tend to be rapid, automatic, and non-consciously directed; these judgments affect cognition, biasing subsequent judgments in the direction of the initial valence in search of congruence both in the moment and over time (hence the wisdom of the saying that “you never get a second chance to make a first impression”) (Lodge and Taber 2013: 20-21). Hot cognition is a key element of motivated reasoning—it reflects a model of the rationalizing individual who, after some initial, affective judgment, works backward to *rationalize* their (often stubborn) judgment.

In this model, deliberation is a tool for the ex post facto rationalization of a judgment, rather than the cause of it, where “affect precedes and contextualizes cognition” (Lodge and Taber 2013: 22). Hot cognition is typically a system 1 cognitive procedure in which affect (i.e. emotion) drives and conditions perception; it stands in contrast to “cold cognition,” which refers to system 2 rational, critical, reflective cognition that is lacking in significant, directly affective content. (As a rough heuristic, think of the Bayesian model of the rational updater as corresponding to system one thinking and, within the context of the elaboration likelihood model, high elaboration). Hot cognition is thus primarily driven by affect while cold cognition is

driven by rational reasoning, with the former tending to rely on heuristic-driven information processing and the latter relying on systematic reasoning (Chaiken and Trope 1999, Kahan 2013). (Affect is likely present in each mode, but it is *far more prevalent. Indeed, it is primary*, in the hot cognition mode). Moreover, while individuals can be cued in some circumstances to engage in Bayesian-style reasoning (Druckman 2012; Taber, Cann, Kucsova 2009), this is not the default or common mode of reasoning. Indeed, prompting such a turn to cold consideration can *backfire* (Kahan 2013).

### *Directional goals, accuracy goals, and reason-giving*

Individual instances of cognition leading to a judgment can be classified along a continuum that stretches from (potentially non-conscious) high directionality imperatives to (potentially non-conscious) high accuracy imperatives. Reasoning at any point along this continuum can be motivated, and more or less reasoning effort might be employed, depending on the individual, their motivation to reason, and the context in which reasoning takes place (Leeper and Slothuus 2014: 140-142). Directional imperatives tend to be lodged in system 1 thinking and are concerned with satisfying affective considerations—especially the maintenance or defense of certain core or high-priority beliefs, attitudes, or elements of one’s identity. Accuracy imperatives are concerned with achieving the “right” or “correct” judgment (i.e. valid judgments that accurately represent the reflective will of the individual) and tend to be a product of reasoning cued by the *motivation* to get a judgment right<sup>49</sup> (the word is used here

---

<sup>49</sup> There may be affective motivations to pursue accuracy (see Toure-Tillery and Fishbach 2014). However, it seems most common (and most problematic for deliberative democracy) for affective considerations to generate directional goals.

not in the sense of biased motivated reasoning but in its more traditional sense of being compelled to do something for some reason). Of note, however, is the fact that more cognitive effort—system 2, high elaboration—does not *necessarily* push individuals towards right, correct, or good judgments; what matters is the *type of cognitive strategy employed*, since increased effort could be deployed in search of achieving directional goals as well as accuracy goals (Kahan 2013, Leeper and Slothuus 2014: 141-142). So, judgments are the outcome of, among other factors, underlying motivations, specific cognitive strategies, and intensity of effort. Thus, when it comes to democratic deliberation, it is not merely enough to suggest that deliberation itself is a cure for motivated reasoning. The particular *approach to* and *design of* a deliberation matters significantly, as does the motivation of the individual who is deliberating to get the judgment right.

Leeper and Slothuus argue that motivated reasoning does not necessarily imply biased reasoning—this point acts as a critique of the model offered by Lodge and Taber discussed in this chapter, since their characterization of motivated reasoning implies that such reasoning is inherently biased—given that the state of motivation has as its object some particular goal; in support of this claim, the authors quote Fishbach and Ferguson (2007), who note that motivation is merely “a cognitive representation of a desired endpoint that impacts evaluations, emotions, and behaviors” (491). Goals are then divided into the same two (broad) categories outlined above: directional goals and accuracy goals (Kunda 1990). The nuance here between these categories is important, given that my concern in this dissertation is not with the existence of motivation per se, but rather *our awareness of it and the degree to which it is made explicit when reasoning in deliberative settings*—especially since the reasons provided by

one individual can generate motivations in another or in others. The challenge of motivated reasoning to autonomous deliberation is thus a problem of generating reasoning strategies at the individual and collective level that are open to critical interrogation by that individual or collective, and which can be honestly and accurately communicated to others in deliberative settings when they are adopted towards some public end (i.e. generating political decisions that will be binding for the collective).

All of this assumes that validity is among the goods preferred as an outcome of individual judgments and decisions, and that has been the fundamental assumption that I have been operating under throughout this dissertation. If the ability to generate validity intersubjectively and translate the products of that procedure into political outcomes is deemed to be unimportant, then it is hard to see why individual participation would be necessary at all, aside from as a perfunctory act of system and policy/legislation legitimization. Democratic deliberation, however, aims at generating better judgments and decisions. One important, but not necessary, function of deliberative democracy is to generate and/or transform preferences through the exchange of preferences and reasons for those preferences; in those cases, validity is a concern for both establishing shared understanding and for producing judgments that lead to shifts in preferences based on shared assessments of the situation at hand. Another function of deliberation is to *clarify* preferences: both for the individual who is deliberating and for others who are participating in that deliberation. Again, though, as we have seen, *more deliberation* (i.e. more cognitive effort towards reaching a reasoned judgment) is *not necessarily better* for producing accurate outcomes: the underlying motivation for reasoning in the first place matters, since different goals will generate different

strategies for reaching a judgment. This is significant because, as Leeper and Slothuus conclude, “with different strategies employed, reasoning toward distinct goals should produce heterogeneity in reasoning outcomes rather than convergence on ‘best’ or ‘correct’ outcomes” (Leeper and Slothuus 2014: 141).<sup>50</sup>

Leeper and Slothuus give two examples of divergent motivations for reaching a judgment: in one, a directionally-motivated (partisan) individual employs a party endorsement as a low-effort heuristic for reaching a judgment on some issue or candidate; in the other, an accuracy-motivated individual relies on “expert”/expert<sup>51</sup> endorsement to reach the same end by low-effort means. However, as the authors note, “...the underlying motivation behind the task will affect what evidence individuals select, how that evidence is evaluated, and what voting decision they ultimately make” (Leeper and Slothuus 2014: 141). In this instance, the system 2-search for information (conscious) is directed by system 1 (non-conscious) motivations. If the directionally-motivated individual were to employ a higher effort approach (e.g. increasing the number of data points), it is still likely that the same outcome would be reached, since the *process* of accumulating data is structurally motivated towards some

---

<sup>50</sup> There is also an important *social element* to democratic deliberation and motivated reasoning. Public deliberation is an inherently social endeavor, and among the sorts of motivations that drive motivated reasoning, concerns about “protecting one’s identity or standing in an affinity group that shares fundamental values” are notable (Kahan 2013: 408; see also Cohen 2003 and Sherman and Cohen 2006, each of which is cited by Kahan). Ultimately, one’s membership in a community and their standing within that community are important (psychological) variables that should be taken into consideration when trying to understand what leads individuals to reason (or not) the way they do.

<sup>51</sup> Ultimately the signaling of expertise is what seems to be essential in generating guidance for low-information individuals, rather than some pre-determined, objective standard. While some “experts” may lose credibility if exposed as unreliable, it is clear that many contemporary charlatans continue to guide low-information judgments regardless of the dubious nature of their expert status.

particular goal (e.g. defending their pre-conceptions) (Leeper and Slothuus 2014: 142). In the case of democratic deliberation, it is possible that an individual with a good reason to pursue a directional goal will bring the same sensibility to the process—probably outside of their awareness.

The empirical work done by Leeper and Slothuus that I have drawn on above is about *politically partisan* individuals. However, in broader contexts generally, and in deliberative contexts specifically, individuals are often partisan in one way or another—whether or not that partisanship manifests itself as party identification. Everyone has commitments and loyalties, some of which may be central to his or her identity (e.g. religion, gender, ethnicity). The underlying logic of motivated reasoning is that one has a deeply held preconception or position that they will strive to maintain or bolster, even in the face of compelling reasons to re-evaluate or jettison that attachment, and even at times when specifically motivated *not to* use those commitments as guides for (biased) reasoning.

The issue of concern for proponents of deliberative democracy as an epistemic theory of how to produce good judgments and decisions is thus what to do about the inherent and pervasive tendencies of individuals to draw upon and be guided by markers (e.g. core beliefs, attitudes, elements of identity) in such a way that reasoning becomes distorted and those who reason come to act as *rationalizing* agents rather than as rational (and autonomous) agents. Here I depart from Leeper and Slothuus's more sanguine conception of motivated reasoning. As noted, they do not necessarily see motivated goals as problematic; they argue that motivation does not necessarily imply bias. However, as seen above in the model of motivated reasoning offered by Lodge and Taber and in data from Kunda, Druckman, and Dunning, motivated

reasoning tends to indicate some degree of bias. This tendency should come as no surprise since strong motivation to achieve a goal—especially when mostly or entirely directed outside of conscious awareness—tends to generate biased reasoning strategies: a kind of cognitive version of the “by any means necessary” strategy.

So, perhaps motivated reasoning does not logically entail imply bias by necessity; however, in practice, it does tend to generate biased approaches to reasoning (e.g. selective information choice, ex post facto rationalizing, information discounting) that undermine autonomy as I have defined it. Moreover, it also compromises the ability of individuals to engage in an intersubjective process of establishing validity through reason giving—a process central to democratic deliberation. This compromised outcome occurs, in part, because *context matters*: politics selects for motivated reasoners, which reflects institutional incentives for politicians, partisans, and those with particular agendas or interests to get what they want from the system. While some might argue that this is merely political pluralism at work, it may nonetheless generate potentially inequitable (and irrational, from the deliberative democratic point of view) political outcomes because, among other reasons, “the flaw in the pluralist heaven is that the heavenly chorus sings with a strong upper-class accent” (Schattschneider 1960: 35).

Leeper and Slothuus defend their position by arguing “when the ‘motivated’ in motivated reasoning takes on the connotation of bias, it evokes an unnecessary normative tone” (Leeper and Slothuus 2014: f.19). This argument provides a chance to pivot and return to some of the normative claims that underlie theories of deliberative democracy—what is expected from those who deliberate and the sorts of outcomes that are desirable. While the

literature in social and political psychology does not itself generate normative justifications, goals or standards, it can and does shed light on the plausibility and potential effects of those justifications, goals, or standards that can be taken up by political theorists (and others)—which is in part what I am doing in this dissertation. So, what can we conclude about some of the key normative elements in light of the literature discussed in this chapter so far?

Recall that some theories of deliberative democracy prefer that individuals act in good faith when deliberating, and all theories require that individuals provide reasons for their preferences and are willing to exchange those reasons with others when attempting to arrive at individual judgments and collective decisions in a particular political context that aims to generate space for open, fair, and honest discussion and debate. Recall also that the underlying assumptions of deliberative democracy as an epistemic theory of democracy—a theory that argues that deliberation *tend to produce good judgments and decisions*—is that individuals are rational and autonomous, at least broadly speaking. The former requirements are inherently normative, since they are about how we ought to “do democracy” together; the latter are empirical claims about human capacities and the extent to which they are or can be practiced in a democratic deliberative setting. Taken together, they offer a competitive alternative to the politics of violence, coercion, or the mere aggregation of preferences.

Thorough and compelling justifications of a range of normative arguments for deliberation have been offered elsewhere (for instance, see Gutmann and Thompson 2004; Parkinson et al. 2013; Warren 2002); my focus is on the justification for the empirical claims of autonomy (and, to some extent, rationality) and why these are necessary and important goals that are undermined by motivated reasoning. As I argued in chapters one and two, the

distinction offered by Kant between *autarchy*—the ability to make one’s own choices—and *autonomy*—which implies that the individual has valid *reasons* for those choices—is essential and necessary for engaging in deliberative democratic decision making.

Politics is present in one form or another whenever decisions that affect the collective must be or are made, and politics ends whenever violence or coercion begins, insofar as violence signals the breakdown of politics. In deliberative democratic politics, the currency in which politics trades is *reasons*. Reasons are exchanged whenever decisions must be made, and the requirement to give reasons is among the costs of doing democratic business, especially for democratic deliberation; indeed, reasons are part of the institutional structure that makes democratic politics possible insofar as legitimate reason giving is a part of trust building, generating legitimacy, and justifying binding judgments that individual actors may dislike, disagree with, or be disadvantaged by. Reason giving in democratic politics both helps protect against the arbitrariness of rulers and serves as one of the key means by which individuals are connected to the institutions of democratic governance when they are asked to or required to participate in collective decision making (e.g. through juries, citizens’ assemblies, and so on). These goods are among the reasons why theorists of deliberation set high standards for reason giving and why I have set the threshold for the epistemic defense of deliberation and the individual and collective capacity for autonomy that underwrites it at a fairly-high level.

As I discussed in chapter three, the exchange of reasons helps enable the generation of intersubjective markers that act as a sort of tally (e.g. deontic scorekeeping) that can contribute to creating democratic goods. When I give you reasons, and you accept them, we communicate in such a way that establishes goods such trust, coordinated action, motivation to carry out

future exchanges, and improved chances of compliance. Overall, reason giving improves outcomes and assists in coordinating and justifying those outcomes within the context of collective democratic action. The underlying assumption for the need for reason giving parallels two core assumptions of deliberative democracy (and plenty of other theories, institutions, jobs, etc.): that individuals are rational and that they are autonomous. Rationality requires that individuals can reason more-or-less accurately about the world (most of the time) and connect facts, values, and norms to preferences in a coherent, stable way. Autonomy requires that, as I noted in chapters one and two, via Christman's (1991) definition of autonomy, individuals are aware of and affirm the most immediate and significant causes of their judgments (i.e. their motivations) in a minimally rational, coherent, and accurate way. Each of these requires that we press Kant's distinction between autarchy and autonomy a bit further to add that individuals need *reasons for their reasons*; more specifically, I mean that individuals need to be able to give a more-or-less *accurate* account of the immediate origins of his or her reasons and their underlying motivations for holding them. If the reasons given by an individual turns out to be arbitrary or flat-out deceptive, then, in the long run, on a large scale, trust in the legitimacy of key democratic practices and institutions might erode (and might erode in the short run, too). There is also a risk to the stability of meaningful democratic governance, not to mention the overall quality of the decisions that are generated.

When individuals engage in motivated reasoning, they risk violating the principles of rationality and autonomy that underlie participatory democratic governance in general and deliberative democracy in particular. This erosion occurs because the motivated reasoner is not interested (consciously or non-consciously) in giving accurate reasons for their judgments, but

rather in achieving some particular, (likely) self-serving, affect-driven goal through the selective (non-conscious) privileging of certain cognitive strategies. Admittedly, the goal that the agent is trying to achieve may itself be perfectly reasonable—indeed, they may even have good, rational reasons for holding that core goal.

As Sniderman et al. (1986) point out, echoing Leeper and Slothuus, rationalization does not necessarily imply deception, since such “reasoning backwards” is done to fill in missing portions (the middle bits) of the reasoning chain (429). However, *there is no way to know this during the process of collective deliberation* if this is indeed what is being done (in fact, the individual herself may not even ultimately know). When this occurs, potentially relevant and important information—data, motivations, reasons—is buried and remains inaccessible to both the individual who is reasoning and those she is reasoning alongside. While this may be fine for many day-to-day decisions, the standard for democratic deliberation is higher, and thus motivated reasoning, which undermines autonomy and buries information that may be required for establishing validity (and other things) in a deliberation, is deeply problematic; rationalization conceals important information, rendering deeply private and hidden what deliberative democracy aims to make public, accessible, and open to transformation. The only way to guard against corrupted goals is by providing rational, autonomously-generated reasons, as best as possible, as far down as possible. Moreover, the best way to do this, as I will argue in part two of this dissertation, is through specific individual practices and deliberative institutional designs that will act as kluges or nudges, directing reasoning processes and ways of arguing in certain directions.

## The model of motivated reasoning in summary

We have seen some of the core concepts related to motivated reasoning and how they are potential threats to rational, autonomous deliberation. Before looking more specifically at how motivated reasoning links with each of the models discussed in chapter three, and how the phenomenon specifically threatens deliberative democracy, let us look briefly but precisely at what a model of the motivated reasoner entails. For this, I will primarily use Lodge and Taber's model but will recall critiques and limits of the model where necessary.

Lodge and Taber's model is driven by seven postulates, a few of which I have explored here in detail, but which the authors summarize neatly

Our theory can be captured by seven central claims: information processing is largely automatic, it is infused with feelings, it is embodied in physiological systems, it is impelled by affect, it is responsive to the environment through online updating processes, and it builds momentum through affect transfer and affective contagion (Lodge and Taber 2013: 34).

Thus, the individual reasoner in Lodge and Taber's model is not one for whom deliberation *causes their preferences, opinions, or vote choices*. Instead, their judgments (and their individual decisions) emerge from the processing of affectively tagged political concepts that are recalled from rapid, largely non-conscious on-line running tally of evaluations which are anchored in early evaluations and which tend to be subsequently systematically biased in the directional valence of those early (often immediate, non-conscious) evaluations based on what

comes to be something akin to classical approach-avoidance behaviour (Lodge and Taber 2013: 42-44, 49-51). Conscious rationales for preferences or judgments are thus, in many cases, largely illusory; instead, an explanation of a preference or judgment often serves as an *ex post facto* rationalization of prior unconscious and affectively driven desires. In the models of motivated reasoning I have discussed here, *feelings tend to come first, prior to conscious judgment* (Lodge and Taber 2013: 48). This “primacy of affect” is what drives its fundamental ‘rationalization over rationale’ element, since affective information processing directs cognitive behaviour *prior to* deliberation.

Lodge and Taber contrast their model of the motivated reasoner and the on-line affective tally with memory-based evaluation models discussed above. In the latter, the *tendency* is for rational considerations to be drawn from present and long term memory when needed and weighed carefully in order to generate a judgment that fits best with the question or concern to which that judgment pertains, updating the long-term store of rational reasons for or against some proposition or issue or candidate as needed.<sup>52</sup> In this model, each new judgment is, in essence, a fresh, unbiased evaluation of the relevant data. In contrast, the on-line model presents the individual as constructing their evaluations *in real-time, on the fly*, drawing from associated feelings surrounding the concept, individual, or issue at hand. A judgment is made, and an individual then automatically and non-consciously adds the current affective evaluation to the prior tally that can be later immediately recalled when required to generate a judgment. Specific information is jettisoned, leaving only (or, at least, primarily) an

---

<sup>52</sup> Not all memory-based reasoning is rational and motivated by accuracy, which is a reminder that system 1/system 2, hot cognition/cold cognition thinking and reasoning are often bound up and in immediate tension with one another.

affective impression to act as a future consideration when a new judgment is required (Lodge and Taber 2013: 49-51). Thus, future judgments are heavily biased and firmly anchored in past affective judgments, generating a kind of affective judgment feedback loop. Rather than each judgment being made anew, it is constructed from a limited range of affectively charged materials available for recall.

Lodge and Taber are careful to note, however, that this tendency has some important, enabling value. They draw on the medieval argument of French philosopher Jean Buridan, who argued that a hungry donkey placed equidistant between two piles of hay would be unable to choose which to feed from, and would thus starve to death, to suggest that hot cognition— affectively driven cognition—solves this problem. In the case of hot cognition, some affective intuition enters the equation and impels us in one direction or another; this, incidentally, is also why we do not tend to get stuck in one place, weighing our options about how to get to the grocery store or when to plan a meeting (Lodge and Taber 2013: 44). As Damasio (1994) has shown, affect plays a significant role in generating a *feeling of what we should do*, eliminating what otherwise might lead to a rational stalemate.

Like the problem of Buridan's donkey, Damasio (1994) tells the story of his patient Elliot. After brain surgery to remove a tumor in his frontal lobe (in the prefrontal cortices), a brain region responsible for, among other things, the regulation of emotion, Elliot had trouble making mundane decisions such as where to eat lunch or when to book an appointment (Damasio 1994: 35-39). Damasio argues "intuition is simply rapid cognition with the required knowledge partially swept under the carpet, all courtesy of emotion and much past practice" (Damasio 1994: xiii; see also Kahneman 2011). He sees emotion as an element of reasoned judgments, or

perhaps it is better put: emotion is a partner to rationality. This conception of affect is consistent with Lodge and Taber's model of motivated reasoning. There is nothing *inherently troubling* with affectively driven judgment. However, when it goes wrong, it has the potentiality of going *very wrong*, becoming surreptitiously embedded in our reasoning and even our institutions. For instance, prejudice, a profound and persistent social pathology, emerges in just such a way. For deliberation, motivated reasoning is a significant threat for exactly this reason: it tends to nudge reason giving (and rationality and autonomy) off the rails. Motivated reasoning is thus especially problematic for deliberation as a theory of epistemically good decisions made by rational, autonomous individuals, particularly so because *we do not always know when this occurs*. It is to this point, in a review of how the cognitive models discussed in chapter three relate to motivated reasoning, that we now turn.

### **Cognitive models, motivated reasoning, and democratic deliberation**

One of the key reasons that motivated reasoning is inconsistent with public trust, legitimacy, and the epistemic and other goods of deliberative democracy is that the origins of the goals, strategies, and reasons that are being pursued by the motivated reasoner may be irrelevant to—or inconsistent with—the (specifically) public and political issues under consideration at a given moment. Moreover, they might undermine the ability of a given community to rely on intertwined chains of reasons as debates, discussions, and deliberations continue to unfold. They might also be inconsistent with the range of preferences or values that are accepted as legitimate considerations in a deliberation. For instance, if racism, sexism, or homophobia are masked as some other, legitimate concern around a policy issue, as in the case where “moral

hazard” is used as a defense of limiting social assistance for particular populations seen as being prone to make use of social services when negative affect—dislike—towards that group is the true motivating factor.

Before moving on, though, an important caveat: as shown above cognitive effort alone *is not always* a good predictor of whether one will take a high-directionality or a high-accuracy approach; more effortful cognition can be biased, too. Equally as important, and perhaps related, is the fact that as Sniderman et al. (1986) point out, there is not necessarily an inferential dichotomy between the “head think” of educated, rational, and cognitive driven political judgment and the “gut think” of irrational and unsophisticated lesser educated, affect-driven individuals, since more sophisticated individuals use both beliefs and feelings when making political judgments (426-427). Accordingly, it might be the case that motivated-reasoning-driven missteps in reasoning and judging are a *general* problem, even if the more politically sophisticated are willing, able, and prepared to employ more tactics to generate stable judgments (Sniderman et al. 1986: 427). The key takeaway here is that the problem of motivated reasoning may be a general one, and so there may not be any a priori way to sort out those who are being driven by motivated reasoning and those who are not. This problem means that addressing the challenges posed to democratic deliberation by motivated reasoning will require *structural changes* to how deliberations are run alongside any reforms to *particular* individual practices (as we will see in the chapter five).

So how then are the four models presented in chapter three related to motivated reasoning and what does this mean for epistemic the defense of deliberative democracy? The four models together, alongside motivated reasoning, give a plausible and worrying account of

the citizen-deliberator. The fundamental emergent concern for epistemic theories of deliberation is the revelation of where judgments might be generated and how they are presented and defended: away from what is traditionally expected and required as the mode of deliberation—the rational, conscious, autonomous mind—to a mode potentially less favourable to deliberation—the irrational, affective, unconscious, externally directed mind that may or may not represent reasons in an accurate way. In the following section, I will link the cognitive models discussed in chapter three with motivated reasoning to show how they link together to create a potential problem for theories of deliberative democracy that rely on claims of epistemic superiority, grounded as they are in the required individual capacities for rationality and autonomy.

As noted, the cognitive mode required for deliberation is one marked by rational, conscious, and autonomous reasoning; these are necessary to maximize the probability that preferences, judgments, and the reasons for them can be accurately communicated from one agent to another so that all parties involved can operate on a shared understanding and generate potential outcomes that are mutually understood, the underlying justifications of which are equally clear and open to interrogation. Returning to the systems model popularized by Daniel Kahneman and Amos Tversky, recall that system 1—automatic, rapid, non-conscious cognition—is responsible for generating preferences and judgments outside of awareness (motivated or not) of system 2—effortful, slow, conscious—considerations. In Petty and Cacioppo’s Elaboration Likelihood Model of Persuasion and Attitude Change, this roughly corresponds with low elaboration—the peripheral route of information processing that is heavily reliant on affect, cues, and heuristics. It is in these modes of cognition that cognitive

distortion is most threatening; indeed, in these modes, one's rational guard is down, increasing the probability that information from and about the world will be warped significantly en route to the generation of a judgment. It is in this mode that system justification (Jost), social intuitionist reasoning (Haidt), and automaticity (Bargh and Chartrand) operate and are rationalized. Indeed, it is instructive and revealing *to think of these three models as explaining (among other things) common and systematic types of motivated reasoning in the mode of system 2/low-elaboration cognition*, which are then, when necessary, rationalized after the fact.

The most troublesome implication for deliberative democracy here is the distortion in reasoning that is generated by motivated reasoning, most commonly presenting in the system 1/low-elaboration cognitive mode but also creeping into system 2 thinking. This distorted cognition threatens both rationality and autonomy, and thus threatens to undermine the ability of those who deliberate to intersubjectively establish validity. Such distortion generates at least two types of challenges: responding to the existence of subterranean motivations and sorting out distorted reasoning chains. To further complicate matters, each of these challenges out at two levels: internally and externally. So, addressing these challenges requires knowing where to look for them.

Before discussing these two sorts of problems, I will explain the two levels at which they operate. The levels correspond to *who* is directly affected by distorted judgments. Internally, the individual is affected insofar as he or she is not self-directing. This lack of self-direction is a failure or shortcoming of personal cognitive autonomy, and to the extent that judgment is distorted outside of the awareness or subsequent control of the individual, he or she is less the

author of his or her life, which has implications discussed in chapters one and two surrounding the value of self-determination. Externally, the group of which the individual is a part is affected by distorted judgment and constrained autonomy insofar as they are unable to rely on the reasons the deliberator is giving. This failure undermines the democratic goods I discussed above and in chapter three—trust, future motivation, preference stabilization, compliance, and so forth.

As for the sorts of problems arising from distorted judgments generated by motivated reasoning, first, subterranean motivations emerge when the individual who is forming or expressing a preference or judgment is unable to account for the true (immediate) causes of that preference or judgment. Again, this challenge emerges as especially threatening to deliberative democracy, since theories of deliberation are founded on, among other requirements, the need for truthful, accessible communication. Subterranean and inscrutable motivations that affect preferences or judgments outside of the awareness violate the principles of rationality (since reasoning is required but not available in such an instance) and autonomy (since an accurate account of reasons cannot be given for the motivations). What is left is at best autarchy—the ability to make decisions for oneself—without autonomy—the ability to give (accurate) reasons for those decisions. What is thus lost for the individual is greater control over her judgments and, perhaps to some extent, her destiny.

Second, distorted reasoning chains emerge in motivated reasoning when affect precedes and directs judgment outside of one's awareness. The individual begins with some particular valence towards the matter at hand and then skips to a corresponding conclusion that satisfies the maintenance of that affective position or some broader psychological need.

After the fact, if required to, the individual might *reason backward towards the middle*, filling in the missing middle parts of the reasoning chain (i.e. the specific reasons or arguments for or against some preference or judgment) (Sniderman et al. 1986). An example will help to illustrate how this might work. Sniderman et al. (1986) studied broken reasoning chains within the context of American policy reasoning regarding government assistance to African Americans. They found that many (white) respondents with lower levels of education began with their like/dislike of the group, jumped to a policy position that reflected more or less approval of government assistance for African Americans, and worked back from there to justify their position so that the explanation given for a policy preference merely a rationalization. Those with higher levels of education were also found to reason backward, though *ideology* drove their rationalizations more often than affect. In both cases, policy preferences, whatever their origins, could be operating as “causes as well as effects” (Sniderman et al. 1986: 419).

In this example, motivated reasoning—to preserve and support initial affect—shapes subsequent “reasoning” (i.e. rationalization), but remains unknown to the subject and thus outside of scrutiny. Of course, the Sniderman et al. study was of survey respondents. In a deliberative context, considerations might be different, as might be the incentives to generate or express certain preferences or judgments. However, as Schulz-Hardt et al. (2000) found, individual biases and the tendency to be motivated to confirm one’s initial views are not necessarily attenuated or eliminated by reasoning in groups; in fact, in many cases when reasoning in a group, the frequency and extent to which bias and motivated reasoning occur increase—especially if the group is homogenous. However, encouragingly, the group

amplification effect is *diminished* when the group in question is heterogeneous (Schulz-Hardt et al. 2000). This implies that, as will be more thoroughly explored in part two of this dissertation, deliberative design might be able to mitigate or transform or even eliminate some of the more troubling cognitive tendencies of deliberators. Nonetheless, the phenomenon of motivated reasoning is common and significant—enough so that any claim by deliberative theorists about the epistemic value of deliberation must remain subject to rigorous evaluation and review.

## **Conclusion and Summary**

In this chapter, I have argued that motivated reasoning—a feature of the models presented in chapter three—poses a *particularly significant* threat to the epistemic benefits of democratic deliberation. I have explained what the phenomenon is and specifically how it may structurally affect deliberation in negative ways that risk undermining the rationality and autonomy of those who deliberate, with reference to hot cognition and distorted reasoning chains. I have, again, argued that while theorists of deliberation should pay careful attention to motivated reasoning, this phenomenon is not itself a reason to abandon democratic deliberation as an approach to generating epistemically good judgments decisions, but, nonetheless, it remains a challenge that will require a great deal of attention is the benefits of deliberation are to be maximized.

## **Chapter 5: Institutional deliberative autonomy and practice-based responses to the psychological challenges of democratic deliberation**

Can we design deliberative practices in such a way that enables them to underwrite and enhance our capacities for autonomy and rationality? That is the question I will answer in this chapter. I root my exploration of this question in the context of our having some good reasons to doubt just how effective our capacities for autonomy and rationality are, as I have argued in the previous four chapters. To answer this question, I engage with theories and empirical evidence from the fields of cognitive, social, and political psychology that have been brought to bear on our concepts of autonomy and rationality throughout this dissertation. Specifically, I will do four things. First, I will outline how institutional deliberative design can address some of the challenges to good epistemic deliberation raised in previous chapters. Second, I will introduce my concept of “institutional deliberative autonomy” (IDA) and propose a new way of thinking about—and a new standard for—conceptualizing and evaluating autonomy and rationality in the context of democratic deliberation. I will further explain how autonomy and rationality might be judged when put into practice as a way of determining whether deliberations generate autonomous judgments and decisions. Third, I will extend my discussion of IDA to examine four concepts—a mix of my own and those of others that I will introduce below—that can be mobilized and applied to the question of how to design deliberation so that it maximizes autonomy and rationality. Fourth, and finally, I will outline the steps that theorists and practitioners should take to move from these concepts to the institution of deliberative practices that maximize the probability that participants will make good political judgments and

decisions. Recall that I define a good political judgment as one that is representative of the preferences of near-fully-informed individuals in ways that are transparent, valid, and reliable; and that I define epistemically good decisions as collective decisions that emerge from good judgments and which also meet the criteria of good judgment.

### **How institutional design can be used to respond to challenges to autonomy and rationality**

One of the foundational assumptions in biology (and other fields including architecture and design, computer science, and international relations) is that structure tends to dictate function (even though regarding, say, evolution, function may also lead structure). The point here is that the way something is set up is not neutral regarding the outcomes it tends to enable or produce. Some structural setups will tend to produce certain sorts of outcomes; others will tend to produce different ones. It follows then that if you change the structure, you have a good chance of also changing the outcomes produced within that structure. Thus, when the goal is to change the direction of outcomes—for instance, towards generating more autonomy or rationality or both—then the question of how to set up particular structures becomes essential one designing a procedure or an institution.

Structures come in different forms. A room is a structure, but so is a set of rules, and so are areas of the brain responsible for certain cognitive or motor functions. Move the windows in a room, and one's furniture arrangement might have to change, as might one's way of using that space. Change the rules of a game, and the sorts of individuals who succeed at that game might change. Damage a certain part of the brain, and you can often expect a different human

being to emerge—or, at least, a different sort of behaviour. In the same way, changing how deliberations are structured—when they are held, where they are held, how speakers are chosen, how and what instructions are given, where information comes from, and so forth—should change the kinds of outcomes (e.g. judgments and decisions) that are produced. It might also change *how* those outcomes are brought about in the first place. This is as true of democratic deliberation as it is of any practice that requires a specific sort of structure in order to operate.

Democratic deliberation as a practice is often highly and intentionally structured in particular ways. The types of decisions just mentioned surrounding a deliberation are significant—far too significant to be left up to chance given that those committed to deliberation are making a *normative commitment* to empowering individual citizens to democratically engage and to take part in a decision-making process. So, when specifically responding to the challenge of a-rational cognition and the cognitive limits that condition epistemically good (or bad) deliberation, those who engage in institutional deliberative design must be aware that some ways of approaching deliberation will serve the particular end of generating autonomous and rational judgments and decisions, and some will not. Indeed, some will work against those aforementioned ends. Designers of democratic deliberation must then take this into account and adjust to maximize the probability that better judgments and decisions—in the epistemic sense—are made.

How? Those in charge of institutional design can respond to the challenges to autonomy and rationality within the context of deliberation by carefully designing *procedures, practices, and even physical structures* in such a way that each responds to specific challenges to

epistemically good deliberation outlined in chapters three and four. This means that specific elements of the practice of deliberation must be tailored to respond to specific challenges related to human cognition, or at least certain *types of challenges*. In this way, institutional design *becomes a bridge between our normative expectations about deliberation and our evolved natural capacities*. In essence, those who work on deliberative design must employ kluges—fixes that work around a problem without eliminating the underlying cause or causes of it (since the underlying causes in this case—our biology, our culture, and even our form of government—are not things we can easily change). This approach, of course, is not a new one; this is not a revolutionary way of navigating the world and our place in it. As Heath (2014) argues, a large part of what has made us successful as a species is our ability to extend our consciousness out into the built environment and to design it in such a way that it becomes favourable to our needs and goals; when we do this, and we often do, design becomes an extension and tool of rationality that allows us to work around our limitations, whether they be biological, cultural, political, social, or some combination of these.

This chapter is concerned with procedures and practices that might extend our rationality (and my implication, our autonomy) in the context of democratic deliberations aimed at producing epistemically good judgments and decisions. It does not address the question of physical space—that is a separate consideration requiring a separate (though important) inquiry that includes a critical understanding of architecture and the biology of perception; instead, I focus on particular setups that are related to practices and rules within deliberations. Of course, to produce a certain kind of outcome, such as epistemically good judgments and decisions, we require that those who design deliberative institutions prioritize

the epistemic benefits of deliberation since some elements of deliberative design will require tradeoffs against others. For instance, focusing on producing good judgments and decisions might require a great deal of time, compensation, and other resources, which might limit when deliberation is appropriate and for which questions it is appropriate as a means of decision making. Under certain circumstances, this might also limit who is eligible for engaging in deliberative processes—which brings us to the question of who can deliberate, who cannot, and under what conditions.

### **Tradeoffs and limits: Who can and who cannot deliberate**

Choosing to make certain political decisions through democratic deliberation implies certain requirements and tradeoffs that will mean that some people will be excluded from deliberation and others will contribute and take away less from it. This is unfortunate, but it is true. Anyone who denies this sacrifices the facts for the comfort of reassuring but unhelpful untruths. For instance, those with severe cognitive disabilities may be unable to participate in deliberations structured around the goal of generating epistemically good judgments and decisions. *The rationalist approach that underlies this kind of deliberation is neither neutral nor universal.*

There are many ways to engage interpersonally in the world, politically or otherwise, and there are other ways to deliberate. Exploring these alternatives is not my project, though working out how to best design and employ those alternative approaches is important—and I would argue, complementary—work. However, this dissertation is a work concerned with epistemology, not ethics, given that my concern is how a certain range of individuals will operate within the constraints of a given system. Whether *the system itself* is desirable, appropriate, or legitimate

is a crucial question—I think it is—but one that is at best tangential to the questions that I am asking and trying to answer in this dissertation.

Still, by choosing a general rationalist approach—as the style of deliberation I am concerned with here does—those unable to engage at a basic level *in the context of the given approaches to judgment and decision making* will be unable to participate themselves (though, in theory, a representative could be chosen to act in place of another). Limits to who can participate and how is, admittedly, one of the costs of this kind of democratic deliberation. However, one of the general arguments that I am advancing in this dissertation is that *epistemically speaking, we can do deliberation better or worse within a given system (itself open to critique and revision) and with reference to the goals of enabling autonomy and rationality; accordingly so we should design deliberation so that we do better while leaving plenty of space for other meaningful sites of political engagement.*

I am not arguing that deliberation is the only way to make decisions or suggesting that it is or should be the only mode of political interaction in contemporary democracies. Other approaches to political decision making and affecting political change include parliaments and protests, town hall assemblies and meetings with elected representatives, civil disobedience, petitions and letter-writing campaigns, elections, *other* forms of deliberation, and so on. *Any given space or set of rules or practice* will have inclusionary and exclusionary elements, regardless of their origins or purpose, though many may be set up in such a way that they allow for maximum inclusivity when appropriate. The structures and practices of democratic deliberation are well known in a general sense, though within them better or worse judgments and decisions might be made, and more or less inclusion might be practiced. And while, again,

this dissertation in general (and this chapter in particular) is concerned with the epistemology of judgment and decision making within deliberative contexts, it also seeks to imagine ways that deliberation can be more inclusive alongside these goals *insofar as such inclusion will also lead to the generation of better epistemic individual judgments and collective decisions*.

While it might well be true that a significant portion of the population will be unlikely or unable (or unwilling) to overcome their biases, it is not clear that this is enough of a reason to exclude them from deliberation since a) they have a pre-existing right under the democratic commitment to equality to be present; b) we cannot know for sure ahead of time who will and who will not be able to overcome their biases and deliberate well; and c) the presence of such individuals might have a salutary effect on the process and/or outcome of a deliberation. My point is not that we ought to exclude citizens from deliberation; for the majority of individuals, deliberation is (theoretically) a possibility and only a minority in particular circumstances will literally be *unable* to deliberate (e.g. due to severe disability or age); my point is rather that while there will always be some individuals unable or unwilling to deliberate, there will be many who choose to engage but take away from/contribute less to deliberation and that fact should be built into our expectations about what deliberation can (and should be designed) to do. This is regrettable, but it reminds us why other parties in a democratic system (e.g. elected representative bodies, advocacy groups, and so forth) are so important.

There are limits to the effectiveness and inclusivity of deliberation other than cognitive ones. While cognitive limitations are *intrinsically exclusionary*, some limitations are circumstantial (and thus perhaps capable of being remedied). These circumstantial limitations on participation include those that impact individuals who do not have the time, money, or

perceived levels of education required to deliberate, who do not wish to engage, or who are unable to physically attend due to geographical challenges or a lack of awareness that deliberations are being carried out. These limitations are somewhat easier to address *theoretically* given that they are about *resources* and not capacity—though some are structural problems, and are plenty tricky on their own. To address them, though, a normative commitment is required on the part of the state and/or bodies who organize democratic deliberations (especially representative bodies). Extensive resources are required to facilitate the engagement of many segments of the population, *especially those who are historically removed from political engagement*. This requires that deliberations are well funded, and participants are sufficiently compensated for their time, held at a reasonable time and that concessions are made for those with special needs. After that, it is up to those who design deliberations to ensure that the educative function of deliberation is maximized and that those who are historically marginalized voices are heard, including, but not limited to, the undereducated, the poor, ethnic minorities, women, and the LGBTQ community.

### **Habermas and autonomy**

In this chapter, I will focus on the *process of generating autonomy as a capacity exercised by individuals in a group setting*. Again, I am concerned with cognitive autonomy (the ability to give *actual* reasons why one is for or against something) and not autonomy as either freedom or license, though the former understanding is in many circumstances dependent on its two latter understandings. It is dependent because context matters when it comes to cognitive autonomy. In this dissertation, I am specifically interested in the context of public deliberation.

A note on Habermas here is helpful for elucidating exactly what kind of context this is and how it fits into both deliberative and broader liberal democratic systems, and what normative force and purpose (better) autonomous democratic deliberation enjoys.

To begin, it is helpful to outline autonomy at the private and public levels in order to establish the broader context in which my research in this chapter is situated. Habermas (1996) understands the difference between liberalism and civic republicanism, in part, as resting on the fact that each has a particular, different understanding of autonomy. The former understands it as freedom from constraint—exercised individually towards the end of private self-determination—while the latter understands it as a collective good exercised by and for the community which seeks to realize its common self-determination (as a political body). While each form of autonomy is “equiprimordial” for Habermas; so, he collapses the division between these conceptions of autonomy in one important way: he argues that public autonomy, which must not fundamentally override essential liberal human rights, requires a robust civil society in which informal (or semi-formal), un- or semi-organized individuals and bodies such as the media, service groups, voluntary organizations, and so forth, engage in a process of preference and will formation. In a well-functioning democratic system—what he calls a “rational” and “justifiable” system—the preferences and wills formed in civil society tend to penetrate the formal political realm where legislatures (or the like, such as cabinets or single executives) make decisions. So, public autonomy implies the free formation of preferences (and will), though it is (sometimes) informally linked to the realm where formal political decisions made.

Private and public autonomy are prerequisites for cognitive autonomy. The former is necessary because it is the site of private self-determination, reflection, and capacity building

(which are essential for both private *and public* goods); the latter is necessary because it is the site of the *political application* of both private autonomy and cognitive autonomy as collective self-determination. After all, the private realm, while not explicitly concerned with politics, is none the less a de facto training ground that prepares one for participation in the public realm, should they choose to participate. Indeed, many of the skills required for public participation are built (or not) in the private sphere (see, for instance, Brady et al. 1995). With these distinctions established, going forward it is important to keep in mind that when I discuss autonomy, I am referring to cognitive autonomy unless otherwise stated (although in this section I will qualify “autonomy” more often and explicitly, since I use both senses of the term often).

Now, one of the central aims of this dissertation is to ask when and how individual (cognitive) autonomy fails and what this means for individuals and groups (via outcomes, i.e. judgments and decisions) who are engaging in structured democratic deliberation in civil society and the public sphere. This is important because the better the outcome of individual judgments, the better the decisions made by a deliberative group, and, I am theorizing, the more effective, justifiable, and rational the communication and outcomes will be in the interactions between civil society and the political realm. The generation, stabilization, and communication of judgments is then not only a good in the civil society, but also in the formal sphere. To the extent that judgments are better than they might have otherwise been, they ought to provide better fodder for formal political decisions.

I share Habermas’s concern with the relationship between civil society and the political sphere. The latter, after all, as Habermas notes, enjoys its legitimacy in part through its

acknowledgment of and its acting upon outcomes from civil society, including democratic deliberations. My primary concern in this chapter is how we can *understand and evaluate the process of generating cognitive autonomy*—and thus better meaning and understanding—through collective democratic deliberation. To the extent that we can better understand and evaluate this process, we can begin to improve it; as we improve it, we can also improve the dialogue between the informal and formal political spheres.

In short: a greater capacity for cognitive autonomy (and rationality) by individuals, when put to work while they deliberate, generates better deliberation, which improves the relationship between civil society and legislative bodies and yields better outcomes (and enhances non-cognitive autonomy). And since democratic deliberation is the site where individuals exercise their individual (cognitive and non-cognitive) autonomy *collectively*, democratic deliberation enjoys a special status and requires special attention because it is an institution that both shapes individual autonomy (both kinds) and affects outcomes in the formal political realm. But it is also true that *better deliberation enables individual (cognitive) autonomy* such deliberation leads to better-informed participants and generates both clearer and more stable preferences through iteration and trust-building (Ostrom 1997). Insofar as democratic deliberation is about clarifying and justifying preferences, (cognitive) autonomy both enables and is enabled by deliberation at both the individual and collective levels. In this chapter I am, however, primarily concerned with how the intersubjective process of deliberation can be understood and measured as a *collective* process of preference generation, clarification, and stabilization despite individual-level cognitive biases, distortions, and limits.

Hence my focus on institutional, collective deliberative autonomy as a measurable, procedural, institutional *corrective* (and not just a washing out of random error through mass participation).

## **Institutional deliberative autonomy**

### *The problem*

As I argued in chapter two, *cognitive autonomy* (as opposed to autonomy as license or freedom) can be understood in terms of its relation to internal individual agency in a given context but it is affected by others and it helps to understand it relationally (see my discussion of Nedelsky and relational autonomy in chapter two). Autonomy in the sense that I understand it for this dissertation is specifically grounded in free choice and rationality, albeit in a constrained sense—but it remains tied to an intersubjective process of reaching mutual understanding and generating meaning (in this case I am specifically focused on the process of democratic deliberation). An autonomous agent is one who is able to choose freely—which implies a range of options; who can consider those options rationally (in the two general senses mentioned in chapter one: both instrumental rationality and practical wisdom) before making their choice—which implies a certain level of cognitive capacity; and who can or would affirm all of the factors that went into his judgment (Christman 1991; Elstub 2008; Habermas 1996; Hurka 1993). Again, this general sense of the concept of autonomy goes back at least to Kant (2012 [1785]), who made a distinction between autarchy, which is the ability to make choices for oneself, and autonomy, which implies the ability to give accurate and meaningful reasons for one's choices. The fact that the *actions of others will have an effect on this process* is very important; but the individual is the ultimate location of the exercise of autonomy (however it is

brought about). Nonetheless, since democratic deliberation is an intersubjective process, even individually-expressed (cognitive) autonomy must be understood relationally.

So far, in this dissertation, I have shown how a-rational stimuli—in various potential manifestations—and distorted cognition have the potential to interrupt and undermine autonomy and rationality by activating goals, considerations, frames, motivations, and moods outside of one’s awareness, by distorting how information is processed and constraining the range of stimuli that is considered as relevant to questions and issues at hand, and by obscuring motivations for holding particular preferences and the reasons that support those preferences. This potential for interrupting and undermining of autonomy and rationality presents the possibility of structural, deleterious effects to democratic deliberation, especially as it is defined in its ideal form as a place where the “unforced force of the better argument” (as Habermas puts it) prevails. I have also argued that in response to this, deliberative theorists could review the regulative deliberative ideal, just as scholars of autonomy could review the ideals of autonomy and rationality. However, such a review—insofar as it would require that we lower our standards—could undermine the very point of an ideal: to set a high yet plausible standard for individual, organizational, or institutional practice.

Rather than abandon our normative expectations about deliberation and deliberators, I am suggesting that what deliberative theory requires in response to the threat of a-rational and distorted cognition to autonomy and rationality is a two-track response. First: we must carry out a conceptual re-evaluation of *what autonomy means* in a democratic deliberative context; and second: we need to develop a specific set of institutional design principles/concepts and personal responses designed to assist in bridging the gap between the ideal of autonomy, the

desired goods and outcomes of deliberation, and the challenges we face given the limits of our cognitive capacities and how they interact with our institutional environment. Before moving on to the specific concepts required for such a re-evaluation, I will outline how in general we can rethink our conception of autonomy and rationality in a deliberative context characterized by the goal of generating epistemically good judgments and decisions.

In her book *Habermas, Kristeva, and Citizenship*, Noelle McAfee (2000) argues that modernity is a project: "...a long inquiry into the possibility of whether human beings can use their reason to be independent, free agents able to create just societies" (7). Such an enterprise, she later notes, requires autonomy: a capacity assumed to be possible and championed as the keystone of modernity by theorists as philosophically diverse as Locke, Kant, Rousseau, Hegel, Marx, Mill, Rawls, Sandel, and Habermas. Theorists of deliberative democracy have taken up the banner of autonomy as well, alongside the long project of modernity, to argue that it is both desirable and possible for humans to mobilize their capacity for autonomy—and other capacities—to bring about more fair and just political outcomes and arrangements (Gutmann and Thompson 1996, 2004; Habermas 1996). However, as we have seen, there are internal (i.e. cognitive) and external (i.e. institutional) challenges to the stability and integrity of individual capacities for autonomy. Some of the very cognitive processes that have made us human—or, rather, that have tended to define behavioural tendencies within the species—and which have allowed us to develop civilizations and to flourish as a species, routinely get in the way of our desire to reach, and to perhaps even extend, the ideal of deliberative political citizenship and democratic deliberation. What then, in light of all of this, does idealized autonomy and rationality look like in the context of deliberative democracy?

Conceptually, this question is fairly easy to address insofar as it asks to describe an ideal state: *deliberative* autonomy would reflect a state in which agents were able to minimize or eliminate any extraneous cognitive processes or stimuli that interfered with their ability to reach a cognitively autonomous, rational, and publicly accessible position on a given issue based on relevant, full, readily available, and accurate information. This sense of autonomy links up well with the definition, via Christman, offered above and used throughout this dissertation. Added to this definition, for the purposes of focusing on deliberative democratic citizenship specifically, is the requirement of relevant, readily available, and accurate information, since in deliberative contexts one is only truly as autonomous as the accuracy of factual information and their ability to understand reasons and perspectives offered by others (and their ability to process it) allows them to be since the aim of the ultimate output of a deliberation is a public one and not just a private concern.

However, the limits of our cognitive capacities and tendencies mean that even with perfect information autonomy can be undermined by internal processes and external distractions that are registered non-consciously, but which nonetheless impacts cognition. So, any conception of deliberative autonomy needs to account and adjust for such challenges in its characterization of *the practice* of deliberation, if not in its setting out of a deliberative ideal. Our intuitive response to this problem may be to turn to the concept of “the *econ*” as our primary understanding and regulative ideal of the individual: the mythical being driven by pure rationality and protected from all but the most rational of ends-means considerations when making a decision. However, as demonstrated above, challenges to autonomy emerge from the architecture of our being: from the very structures and processes that make us human and

allow us to function in a complex environment, even if the social, political, cultural, and technological structures that we have developed outpace in certain ways our ability to process life within them. So, rather than attempting to double-down on a deeply flawed and often unhelpful concept of human nature—one that misunderstands rationality and emotion—we need to find ways to leverage the limits of cognitive architecture for the purpose of bringing about more effective, productive, open, fair, and legitimate deliberation.<sup>53</sup> Nonetheless, as noted, there is some value to this concept and some cases in which such a conception of rationality matches human behaviour (see, for instance, Van der Straeten et al. 2010). However, it is not the *only way*, and indeed not even the *primary way*, that humans engage in the world.

### *The solution*

The concept of deliberative autonomy I will pursue here is aimed at understanding and addressing the challenge of individual autonomy by kicking the problem up to the collective level—since that is where each capacity for autonomy is exercised and because the process of deliberation will affect this capacity (for better or worse). It is meant to set up a deliberative ideal based on collective, intersubjective autonomy: what I am calling “institutional deliberative autonomy” (IDA). This concept is grounded in two major claims: one claim is about the normative ideal of deliberative citizenship as outlined above, and the other claim is about the empirical (and demonstrable) limits of our cognitive capacities—specifically as these limits

---

<sup>53</sup> Insofar as the econ is a regulative ideal, it is occasionally appropriate (though we can debate how appropriate it is as an ideal). It is when the ideal becomes an account of how humans *actually behave* instead of how some think they *ought to behave* that things go particularly pear shaped.

relate to the possibility of autonomy. IDA is related to four conceptual elements (each explored below in the next section), with some developed by me and others taken from scholars of deliberation of decision making, directed towards implementation at the institutional deliberative level and aimed at designing institutional channels and outlets to attenuate or transform cognitive limitations in deliberative contexts into desirable outcomes. However, each concept is meant to fit within the conceptual re-evaluation of autonomy that I am carrying out here.

Rather than being grounded in a process focused solely on the individual, *institutional deliberative autonomy is set at the institutional level* so that it can be cast as a collective state and linked to an *evaluation of the procedure of judgment and decision-making among individuals with a common goal: generating shared understanding or meaning*. As noted, traditionally, conceptions of autonomy have been grounded in the individual—and that makes a great deal of sense since our experience of the world is premised on the idea that I am me, you are you, and each of us has a distinct subjectivity over which we have control. However, as we have seen, our autonomy is regularly compromised—in large part because of interactions with others (though not exclusively). And while it nonetheless makes sense to maintain an individualized understanding of autonomy for much of our day-to-day lives, when it comes to deliberation and the desire for epistemically good judgments and decisions, then a specific, group understanding of autonomy is both sensible and useful since the ultimate aim of democratic deliberation is to produce a certain kind of public outcome and not just a private one. Recall that while a capacity for rationality or autonomy *resides in the individual*, it is commonly shaped, practiced, enabled, or constrained alongside and among others, making

these capacities fundamentally intersubjective. The fact that these capacities are conditioned is especially important to keep in mind when discussing democratic deliberation in general, and the epistemic goals, procedures, strengths, weaknesses, and limits of such a deliberation. So, IDA is a capacity exercised individually but achieved collectively through shared institutional arrangements and practices (discussed below); the question, then, is how IDA can be approached or fully achieved—and how we can identify or evaluate such autonomy when it is present.

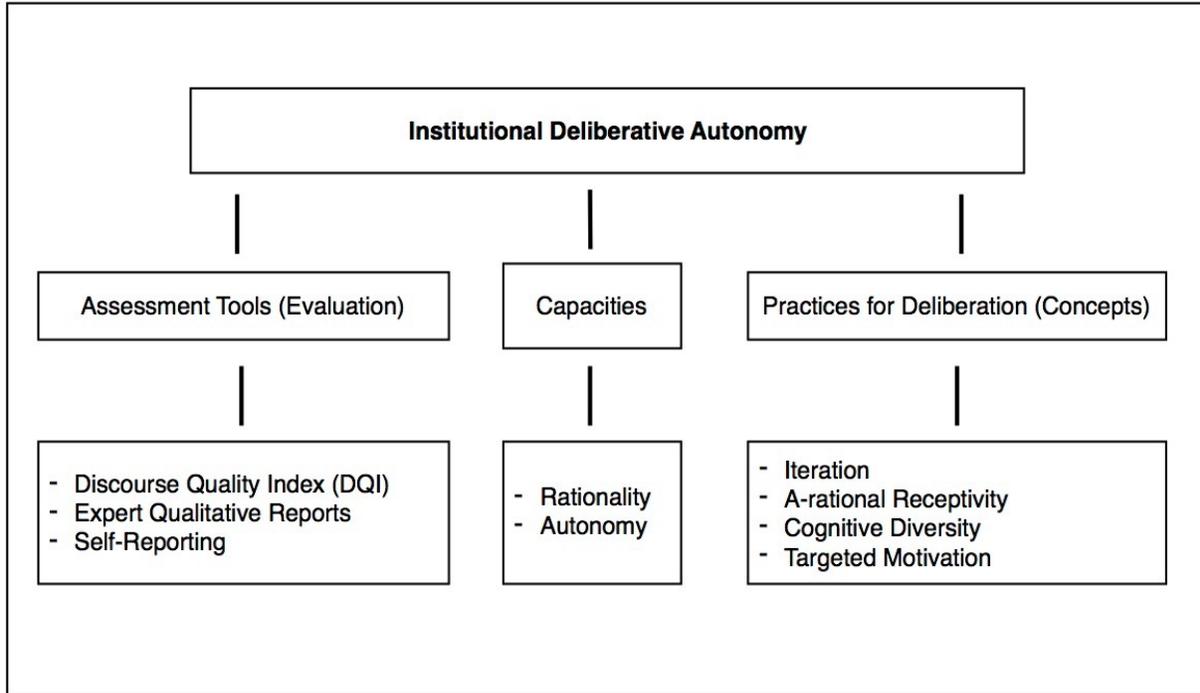
One reason why IDA makes sense as an ideal state that can be evaluated through the quality of a collective process of epistemic deliberative output is that deliberative citizenship is *necessarily bound up in an attempt to establish collective understanding or meaning*, which is a necessarily *intersubjective* experience. The experience is intersubjective because, first of all, deliberation is *formative* of preferences rather than *aggregative* (i.e. a mere aggregation of existing preferences). While most common understandings of politics assume that preferences are pre-formed and individually-located, and so politics is about mobilization or defending preferences, theories of deliberation assume that, at least to some degree, participants will form or reform their preferences through exchanges with others. That makes deliberation an inherently dependent and cooperative endeavour undertaken by a group of people. For another, those who deliberate are often asked to reach a decision that represents the interests of some actors while being accepted as legitimate by all—though still open to future debate, discussion, amendment, or revocation.

These goals are not employed with an aim to undermining the educative function of deliberation: a function which includes both building individual capacity and transferring

relevant information to citizens; instead, the collective procedural goals of deliberation are necessary structural elements for managing a diverse population marked by procedural and moral disagreement: yet another reason for the need to understand autonomy as a collective action. *Accordingly, I define institutional deliberative autonomy as a state of intersubjectively dependent individual autonomy supported by institutional arrangements and procedures designed to maximize individual autonomy and rationality towards the end of reaching shared understanding and meaning based on full, public, and accessible, factual information and reasons.* This definition is meant to capture both the idea of individual autonomy required for good democratic deliberation *and* the fact that approaching (or reaching) this ideal is an inherently collective and institutional affair; it is also meant to highlight that IDA is a state, but that can be evaluated in ways that allow us to then develop approaches to produce or enhance it in future iterations.

IDA includes and is understood through a series of components: assessment tools to evaluate the capacity-status of those who deliberate and to produce IDA (the assessment component), conceptions of core capacities (i.e. rationality and autonomy) that underwrite IDA (the capacity component), and prescribed practices for developing IDA (the practices component). These three components fit together and can be broken down further into sub-components and areas of concern. Figure 1 maps out these components, each of which feeds into IDA.

Figure 1. Institutional deliberative autonomy and its components



This definition of IDA assumes that the ultimate output of a deliberation is one that expresses collective (individual) autonomy through a well-designed process that is designed to take into account that while the individual is the specific site of the exercise of cognitive autonomy, the process of reaching collective autonomy and thus shared understanding and meaning is a necessarily intersubjective one. The goal of approximating what each would have ultimately chosen from a state of full information and awareness of the central and peripheral factors that determined his or her choice remains: especially a-rational cognition, including emotions, feelings, and moods. However, the ideal is now set at the group level and takes into account the *collective, intersubjective* nature of the process of deliberation, which functions as a way of generating autonomy through developing, clarifying, and stabilizing reasons and preferences over time.

I conceive of IDA being evaluated along a continuum depending on how closely collective outcomes match the ideal of a (collective) decision reached in which each member, armed with all the relevant information for an issue, is aware of what is motivating their decision while they are engaging in the intersubjective process of generating shared understanding and meaning. Within this system, to the extent that individual autonomy is compromised, so is the degree of IDA achieved since each affects the collective. And since IDA is an outcome whose quality is (theoretically) located along a scale from more to less autonomous, evaluated against the ideal, we can generate conceptual approaches to improving autonomy, followed by hypotheses for deliberative practices that can be tested in an attempt to improve deliberations so that each deliberative project maximizes its chance to fall closer to the more autonomous end of the scale. (I will provide some specific examples of such conceptual measures in the following section.) This ideal is useful because it provides a conceptual tool that can be used to evaluate performance and, complementarily, to develop other concepts and tools to improve democratic deliberation. It provides us with a target to aim for—one that is normatively grounded, but which includes standards for empirical evaluation that can be based on other concerns (e.g. functionality)

The concept of IDA builds from and extends Habermas' (2003 [1971]) notion of the ideal speech situation, though it differs in two important ways. First, while IDA relies on a (conceptual) procedural measure, it produces a *hypothetical evaluation* of an outcome measured against the ideal of individual autonomy amplified by the composition and desired ends of a given deliberative body. Second, while Habermas notes that individuals in the ideal speech situation shall be free to express their "attitudes, desires, and needs," he leaves

unspecified the role of apolitical and a-rational motivating forces in coming to these conclusions (and fairly so, since that was not his project). In contrast, IDA specifically creates a space for understanding the role of a-rational cognition and cognition distortion in coming to judgments via the modulation of desires, values, and preferences.

It is important to highlight further the fact that IDA is supported by *procedural and evaluative evaluations* based on the ideal of autonomy: as a concept, it includes attempts to capture what the result of what discursive outcomes, individual judgments, and collective decisions would look like when run through an intersubjective procedure—democratic deliberation—that enhance individual autonomy while producing collectively autonomous outcomes. Institutional deliberative autonomy is *institutional* because it is concerned with the rules of the game vis-à-vis deliberation. It is *deliberative* because it is focused on the give and take of reasons and the formation, rather than the aggregation, of preferences—especially vis-à-vis evaluating the validity of outcomes from a democratic deliberation. Moreover, it is concerned with *autonomy* as a capacity that will affect outcomes. At its core, IDA is a concept that comprises an ideal *capacity for autonomy* bound up also in a conception of rationality; it is supported by a set of practices for pursuing that ideal and a series of assessment tools for evaluating the extent to which a democratic deliberative space is conducive to generating decisions suited to that ideal (see figure 1).

## **How to evaluate IDA**

As noted, the ideal of maximized institutional deliberative autonomy sets an ideal standard against which groups of citizens can measure outcomes while accounting for what is often

considered mere background noise to rational decisions: a-rational cognitive processes. In light of this ideal, institutional designs can be mobilized in an attempt to improve deliberations. The assumption is that more autonomy increases the probability that judgments and decisions are rational and epistemically reliable and good. Indeed, this assumption is what animates much of this dissertation and research program.

Of course, the obvious and essential question is: *How can you know if a decision is more autonomous or measure the presence of more or less autonomy, especially if cognitive distortions occur outside of the individual's awareness?* This question is likely the most challenging query for those committed to improving the epistemic value of deliberation through capacity building. There is no standard measure for cognitive autonomy. Moreover, cognitive autonomy is not necessarily a binary state—there are degrees of autonomy that correspond to how much an individual's judgment is formed by factors outside of their awareness and control. It seems unlikely, probably impossible, for an individual to generate judgments of which she has *complete* awareness of the factors that went into it. Indeed, even if she were aware of immediate, first-order causes (e.g. those that were directly relevant to the issue, those that immediately conditioned directly relevant concerns, immediate stimuli that affected her course of reasoning), there is no guarantee that all second-order causes (e.g. past experiences that may be unavailable to conscious memory, stimuli so subtle that it is not registered at all) can ever be fully known and understood.

Because of this, IDA is supported by a series of rough, aggregate, conceptual evaluative mechanisms that seek to provide the normative foundation from which one can evaluate the potential total autonomy of those engaged in deliberation. The range of autonomy runs, again,

roughly, from not autonomous (i.e. fully unaware of the factors that went into a judgment, inconsistent, and irrational) to fully autonomous (i.e. fully aware of all relevant factors that went into producing a judgment, fully consistent, and fully rational) through an evaluation of several procedural measures, which I will discuss in the next section. Each of the extreme states—from not autonomous to fully autonomous—are, of course, theoretical. It is unlikely that either will ever be found in a deliberative context (or any context at all, for that matter). However, they are useful limiting types since they set the terms for measuring *better or worse* outcomes from the standpoint of generating autonomous, rational judgments.

Nonetheless, the question of how to evaluate autonomy remains. Here I want to move away from the word “measure” and instead use the word “evaluate.” The reasoning behind this move is simple: “measure” connotes a more or less exact, quantitative act, like measuring the length of something or aggregating survey data to measure public opinion. IDA, however, is a concept that lends itself to a mix of qualitative and quantitative evaluation and will always, at best, give an approximate idea of how autonomous a deliberation is. One could, perhaps, say that it is a “rough measurement,” but the connotation that comes with the word “evaluate” is much better suited to the spirit of what IDA is theoretically capable of achieving given that its precision will not match that of many quantitative measurements. While the conceptual tools discussed in the next section may themselves be highly measurable (e.g. the Discourse Quality Index—DQI), when added together to examine the degree to which deliberative events promote autonomy, it is more an evaluation than a measurement that occurs.

Evaluating autonomy is still mostly, though not entirely, a much more qualitative than quantitative practice, though it may require a careful balance of the two. Since this is a new

concept, and since it is designed to address a tricky, persistent, and elusive challenge, the evaluative mechanisms of IDA may change over time based on experiences in practice. Some mechanisms might work well, some less well; some will work in some contexts and not in others. The only way to know what works in evaluating autonomy, and when it works, is practice and time. However, we must start somewhere. To that end, I have included three assessment tools for evaluating how well deliberations generate autonomous judgments and decisions at the aggregate level. Each is roughly based on the idea of observable implications—what we should expect to see (or not see) if something is working (or not working). Again, these are subjective evaluative assessment tools—except for the DQI—but as qualitative measures of the effects of deliberation, they ought to help draw general, consistent conclusions about how democratic deliberation unfolds under varying circumstances. *When taken together*, these assessment tools should be sufficient for generating a strong sense of whether more or less autonomous and epistemically good judgments and decisions are being made through democratic deliberation.

Before proceeding, a caveat is in order. Again, IDA is supported by a series of evaluative tools and a set of institutional design concepts, and this also includes, in a sense, a calibration tool. It might not always be appropriate to use IDA's evaluative assessment tools—as some are somewhat intrusive, or, at least, resource-intensive and involved. However, if IDA evaluations are taken for some deliberations and arrangements are determined that generate high levels of autonomy and epistemically good judgments and decisions, then those deliberative models can be transferred to other settings, perhaps without the need to engage in evaluation. To put it plainly: as an evaluative tool, certain procedures may not always be appropriate or required.

Nonetheless, it should be used to help ascertain, as a baseline, when deliberations encourage more or less autonomy.

### *Assessment one: Discourse Quality Index*

The Discourse Quality Index (DQI) was developed by Marco Steenbergen, André Bächtiger, Markus Spörndli, and Jürg Steiner to “serve as a quantitative measure of discourse in deliberation” (Steenbergen et al. 2003: 21). Steenbergen et al. ground the DQI in Habermas’s discourse ethics and code individual speeches on five metrics: participation (i.e. ability to freely participate), level of justification (i.e. whether demands are well justified), content of justification (i.e. whether justifications are rooted in appeals to the common good or individual/group interests), respect (i.e. the speaker’s disposition towards the arguments of others), and constructive politics (i.e. whether the speaker aims at reaching reconciliation or compromise, or generating consensus through their speech).

An assessment of discourse quality includes an index score based on indicators assigned to each of these metrics. (The higher the score, the higher the quality of discourse.)

Steenbergen et al. find that the DQI serves as both a valid and reliable measure of discourse quality grounded in deliberative discourse ethics as they are theorized by Habermas. Such scores are helpful to researchers for determining the extent to which deliberations are unfolding according to broadly—though not universally accepted—deliberative norms. These norms, I have argued, are likely to be conducive to generating more autonomous deliberation—and so measuring them under the aegis of IDA will help researchers ascertain the relationship between elements of a deliberation and autonomous cognition and expression.

According to Steenbergen et al., the DQI is useful for both predicting substantive policy outputs from democratic deliberation and acting as a dependent variable used to explain variance in the quality of deliberative discourse based on the particular context, including the sort and state of relevant institutions in which discourse takes place (2003: 42). Both uses of the DQI are helpful for IDA since index scores can be tracked alongside outcomes from other assessment tools and can inform researchers about whether the quality of discourse co-varies with other relevant outcomes (e.g. good policy, particular institutional arrangements, and any of the assessments that I will discuss below). Most importantly, using the DQI alongside the assessment mechanisms that I discuss in this section could indicate to researchers and practitioners if individual or collective autonomy is associated with high (or low) quality discourse.

It is important to note that while the DQI enhances our understanding of when and how IDA is reached, it might also be enhanced by it. As Steenbergen et al. note, there are four limitations to the DQI: an inability to observe and quantify the authenticity of claims (Habermas regards authenticity as an important part of discourse ethics), the lack of a consensus around Habermasian discourse ethics as the proper framework for evaluating deliberation, the limiting of the DQI to discursive texts—DQI measurements cannot pick up body language, tone, or other non-verbal cues and elements of speech—and the failure of the DQI to capture inequities and inequalities that precede access to deliberation (2003: 43-44). Some of the qualitative assessment tools I present here as part of evaluating IDA may complement the DQI and serve, at least partially, as a remedy to some of these shortcomings.

### *Assessment two: expert qualitative reports*

The following assessment is similar to the DQI, and, indeed, complementary to it; however, this concept is different since its concern is with qualitative evaluation of indicators related to specific approaches to *reasoning and rationalizing* adopted to attempt to track autonomous versus heteronomous cognition. Expert reports would be summaries made before, during, and after deliberations that track proceedings at both the individual and aggregate levels. (Of course, if individuals were uncomfortable with this, numbers could be used instead of proper names for the reports.) Experts can be trained to observe argument styles and to engage in interviews with those who deliberate to ascertain how they are reasoning—or rationalizing—throughout the process. (This could be done by facilitators or by other individuals). In essence, expert reports are a layer on top of self-reports aimed at filling in gaps and making sense of the proceedings at both the individual and aggregate level. As an evaluation of IDA, expert reports would include similar observables to self-reports and the measurement would distinguish itself from the DQI by focusing specifically on the *process of argumentation* as it relates to *autonomy and reasoning*, including, importantly, elements of argumentation that have nothing to do with speech per se.

For one, high levels of autonomy at the aggregate level would imply that the goods from self-reports are observable and more or less confirmable at the expert level. Thus, the observables from the self-reports are directly transferable to the expert level; so, experts might have to pay close attention to not just what individuals are saying, but to their mannerisms, tone, and timbre, moods, and so forth. This approach to evaluation complements the DQI, which, as noted, does not pick up on non-discursive elements of deliberation. As an approach, it

requires further elaboration, however, since we require ways to measure or evaluate such things as tone, timbre, and mood. That is a task for another project.

For another, expert reports should be able to determine whether, over time, the judgments of participants are based on plausible reasons and accounts of their reasoning in the development of their judgments (again, the DQI is a very useful complementary assessment tool here)—especially when a broader range of concerns, including mood, the weather, location, time of day, and so on are taken into account. This requires that experts pay extremely close attention to the deliberation and to those who deliberate. At the very least, extreme examples of a lack of autonomy should be obvious, even intuitively so. Building on research carried out on the role of expert facilitators in deliberation, we have reason to believe that similar training and expertise would make expert evaluation a useful addition to the evaluation of democratic deliberation as it relates to reasoning and autonomy (Mansbridge 2006 et al. 2006, Moore 2012).

There are, currently, approaches to evaluating deliberation similar to what I am proposing. For instance, Gerber et al. (2016) employ an "updated version" of the DQI that includes reporting metrics for reciprocity and equality (see also Gerber 2014). Importantly, Wyss et al. (2015) adapt a specific measure from psychology, "cognitive complexity," that is designed to pick up on the epistemic dimension of deliberation (and also the "social-integrative" element). Cognitive complexity "measures the degree to which an individual perceives, distinguishes and integrates topical dimensions" (2015: 3). The measure picks up the degree to which a participant can approach the epistemic deliberative ideal; as Wyss et al. (2015) note "CC represents an important marker of the epistemic quality of debate; by the

same token, it also implies a willingness of actors to integrate and accommodate other viewpoints and strive for agreement" (2). Expert evaluations such as those that measure cognitive complexity are excellent tools for generating reports on whether or not instances of deliberation approximate the ideals set forth by deliberative theory. Specifically, they can produce a top-level evaluation that participants may not be able to self-report accurately.

### *Assessment three: Self-reporting*

One of the core problems with self-reporting, specifically as it relates to the challenge of a-rational cognition and cognitive distortion to autonomous deliberation, is that if the agent is *unaware* of the effect of certain stimuli and processes on their reasoning, then that crucial information is likely to be left out of the self-report. This problem serves as an important caveat, and it gives us good reason to be sceptical about self-reports. However, there is still a place for self-reporting when it comes to judging autonomy at the aggregate level. For one, we may not know all the factors that go into our reasoning, but that does not imply that we do not know any of them. Moreover, self-reports are an effective element of a broader strategy for ascertaining whether judgments and decisions are more or less autonomous since they generate data that can be evaluated against other mechanisms—as I will suggest below.

Self-reports before, during, and after the full process of deliberation provide a valuable series of data for evaluation. The focus of these evaluations could be broad, but for evaluating autonomy, they would need to include questions about how one specifically came to the judgments they reached and how they feel about the decisions generated. While we may have reason to doubt the quality of self-reports, they are useful nonetheless at establishing a starting

point for researchers from which to evaluate. Moreover, self-reports offer researchers something that they cannot otherwise obtain—an account directly from the individual who is doing the reasoning in a given circumstance.

These evaluations—at least those during and after deliberation—could also include commentary on other participants. Whether questions were open-ended or not, or mixed, the goal of self-reporting is to make explicit the *cognitive process by which judgments and decisions are reached*. Accounts could include, and indeed should include, all known or expected factors, including those traditionally deemed “irrational” or “irrelevant.” This last requirement is essential since the measure of autonomy is not merely concerned with what kind of factors impact a decision—though this matters—but whether participants in the deliberation are *aware of and affirm, or would have affirmed* those factors.

If a high level of aggregate autonomy is present in a deliberation, we would expect to see a tendency in self-reports towards coherent narratives—on balance, though some may arrive at the same conclusions through different approaches—even if there are a range of types of reasons, changes in reasons, and elements of the narrative that are inconsistent. Indeed, *some* inconsistencies or changes will point towards a critical engagement with contrary reasons and arguments, though over time we should expect a clear central tendency to emerge, at least when the available information remains consistent. We should also ultimately expect to see clear reasons why an individual has come to one conclusion over another, which includes reasons why they do not support alternatives. Finally, we should expect to see individuals giving a rough, general outline of the proceedings to date, including an account of what the issue is, why it matters, and how others have engaged in deliberations. Of course, part of the challenge

of autonomy is that motivations and reasons are often hidden from the individual themselves. Self-reports are a first step in evaluating autonomy; a related second step aims to mitigate some of the effects of obscured motivations and rationalizing over reasoning. This requires carefully constructed approaches to deliberative design aimed at addressing these challenges specifically—for instance, the sort of approaches I begin to discuss in the following section after giving a brief summary of IDA.

Nonetheless, we ought to remain sceptical about the degree to which self-reporting produces reliable accounts of reasoning. As Gerber et al. (2016) note, self-reporting often results in social desirability biases that skew results and, when it comes to predicting changes in opinion, do not produce any effect (although external measures do). Again, while we ought to be careful with relying *exclusively* on or even *primarily* on self-reporting, employing self-reports as one tool among others allows for a baseline against which other measures can be compared. More to the point, if changes to deliberative procedures are introduced into a deliberation and self-reports begin to closely match external evaluations, researchers will have a useful proxy measure for deliberative quality—as self-reports improve, we ought to be able to conclude that individuals are more open, self-aware, honest, and autonomous.

### **Summary of institutional deliberative autonomy**

If it is a fair assumption that more autonomy and rationality in deliberations will lead to better judgments and decisions—that is, those that more closely reflect external factual realities and the plausible, full-information priorities of participants—then it is reasonable to watch for changes in the quality of outcomes in deliberations and then to infer from them that something

about the deliberative set up is working. Admittedly, this is a tricky mechanism, and one can never be sure that aggregate outcomes within the IDA matrix *are indeed caused by that particular institutional design and the increased levels of autonomy and rationality they are purported to generate*. Yet, if it walks and quacks like a duck, then it is fair to assume that it is more likely to be a duck than a tank.

Indeed, if better judgments and decisions are being made—more rational and autonomous—on aggregate, then there are some things we should expect to see. At the most basic level, statements should tend to reflect factual realities that are generally agreed on, and there should be a tendency to connect them to reasons, as evidence, and then to specific judgments. Later, these should build towards a decision that plausibly reflects basic facts and the arguments that support them. Even if there is deep and persistent disagreement about the normative implications of certain facts, autonomous, rational deliberation should reveal these and allow for an epistemic process by which a majority of participants can come to agree on an interpretation (which, nonetheless, according to the basic premises of democratic deliberation, remain open to future review and contestation).

Institutional deliberative autonomy is a conceptual ideal supported by recommended assessment tools aimed at discovering the extent to which a particular deliberation or series of deliberations encourages autonomous cognition and a program for generating more autonomous judgments and decisions through careful deliberative design. Towards that end, I have developed (concepts that are my own) and assembled (from other scholars) four mechanisms aimed at moving outcomes towards this ideal: iteration, a-rational receptivity, cognitive diversity, and targeted motivation. Again, these are a mix of concepts drawn from the

work of other theorists—as I note throughout—and some of my own. Together, they form the early conceptual skeleton of IDA insofar as it is an ideal supported by conceptual program for generating better deliberation. This, incidentally, is why these mechanisms are also (broadly) part of the *evaluative element* of IDA: since how well participants perform on them will have an impact on how autonomous deliberations can be said to be carried out.

### **Concepts for deliberative design within the context of institutional deliberative autonomy**

There are many elements of design that must be taken into account when designing democratic deliberations. As noted, some have to do with accessibility and ensuring that wide ranges of voices are heard. Others have to do with assuring a high-degree of likelihood that some decision will be made. My concern is with designing deliberations in such a way that maximize autonomy and rationality so that epistemically good outcomes can be generated. To this end, the four concepts offered below are designed to encourage autonomous, rational deliberation.

#### *A note on facilitation*

The proper facilitation of deliberative democratic meetings by individuals trained in facilitating is essential for the following conceptual approaches to improving deliberative autonomy to be effective. Facilitation could, in theory, be included as one of these concepts. However, I conceive of it more as a distinct and prior tool that enables others. While the subject of facilitation is worthy of a dissertation of its own, I will not address it here other than to say that for the following concepts to be most useful, facilitators will be important to the application of

the practices that emerge from the concepts that follow in this section. It will be up to them to ensure that any concepts that are applied are so done in a way that enables capacities for autonomy (and rationality), among other concerns, to flourish.

As Quick and Sandfort (2014) note “The ability to deliberate does not exist inherently,” which is why facilitators are required. While facilitators are generally concerned with overseeing deliberation, there are several approaches that they can take and some tools they can use to do so. For instance, facilitators can employ votes and/or a collective group statement, push for consensus (or not), draw out participant reasoning, and ensure that alternative views are registered (see O'Docherty et al. 2013). Depending on the purpose of the deliberation, facilitators can engage more or less in steering discussions and pushing back on statements made by participants; importantly, facilitation can be customized for each deliberation based on the ends or outputs desired. Nonetheless, a few, standard types of skills are required for high-quality facilitation.

Quick and Sandfort (2014) draw on facilitation research to identify “key tasks” for facilitators that include (and I quote them, format and all, directly):

- Selecting the processes best suited for accomplishing the task at hand, combining prior planning with improvisation to respond to emerging dynamics.
- Establishing and enforcing ground rules and group norms, particularly maintaining a respectful, open and inclusive environment.
- Supporting diverse participation and manage potential problems of exclusion, power and associated conflict.

- Helping the group work toward its objectives, in part by focusing on relevant topics and managing time.
- Enhancing the development of mutual understanding, for example, through asking clarifying questions, rephrasing statements and supporting diverse perspectives (3).

Accordingly, as they note, the skill-training undertaken by facilitators are aimed at carrying out these tasks. That training includes not only courses and workshops but also practice through repeated facilitations of actual deliberations. Over time, facilitators become an expert just as any professional might through repeated practice. Accordingly, this chapter includes both assessments (above) and techniques (below) that might be picked up by facilitators to focus on assessing and building (high-quality) deliberation through focusing on (among other things) deliberative autonomy.

### *Iteration*

While certain elements of a-rational or distorted cognition may lead to structurally biased outcomes, others are stochastic and thus unpredictable *both within an individual and among members of a group*. The unpredictability of such outcomes implies that, potentially, outcomes at the level of individual choice may be influenced by irrelevant factors that vary from instance to instance and moment to moment, and that are thus highly susceptible to fluctuations due to localized changes. If this is indeed true, autonomy is undermined in the agent insofar as the choices they make and the reasons they give may be unstable. If the effect spreads—say,

through affective contagion (Barsade 2002) or mimicry (Chartrand and Bargh 1999)—then the stochastic nature of the phenomenon becomes, at least for a time, structural, and group outcomes (both individual judgments and collective decisions) may become compromised. If the effects are spread from group to group, it is reasonable to expect that group-level activity becomes the stochastic incident that is part of structural shift *between groups* if there is some occasion for inter-group mimicry through a transfer of that error. This could, in theory, cascade if the effect is repeated in a chain, generating less rational and autonomous collective decisions as it goes.

For instance, imagine a group tasked with deciding between several options for welfare policy—an emotional and controversial issue for many. Imagine breakout groups of four or five people segmented from a larger group of fifty deliberators. Imagine further that there are three or four charismatic individuals who are strongly opposed to, say, increasing rates, and disregard documented decline in living standards of welfare recipients due to inflation, evidence that increasing rates does not lead to moral hazard, and so on. These individuals are personable, affable, even, and persuasive, if factually mistaken about welfare. If these individuals are scattered throughout the group, they may be able to bring fence-sitters to their side based on the force of their personality alone, enabled by affective contagion, mimicry, and other mechanisms. The distribution of these individuals within the group and among the breakout groups was, initially, random. However, their presence can start a cascade that brings others onside—others who might adopt their perspective without having good (i.e. rational, evidence-based) reasons for doing so. Some of these new joiners may then become “carriers” of the message of the charismatic, affable anti-welfarists.

One way to address such a challenge to rational deliberation is iteration: repeated exercises, similarly structured, over the course of a period of time. Iteration is designed to, eventually, develop a stable central tendency, and thus to “wash out” incidences of randomness that might affect particular judgments or decisions (Ostrom 1990, 1997). Thus, deliberative sessions can be repeated—either in person or remotely—several times over the course of days, weeks, or even months with similar issues and questions aimed at generating choices that can be compiled and reviewed (Mansbridge and Warren 2013). (This builds on an old principle that is often related to sports: if you want to know how good you *really are* at some particular activity, do it many in a row, until a mean emerges. If you have ever golfed, this idea should be compelling. After many iterations I learned, for instance, that I am consistently terrible.) Iteration allows for a repeated process that can work against cascading error by providing counter-arguments and counter-frames that will complicate and potentially reverse the deleterious effects of certain mechanisms like mimicry and affective contagion; this is possible since the sorts of preferences that emerge from these affectively-driven interactions may not be lasting. Iteration gives time for such errors to be worked out and for more rational preferences to emerge and stabilize; and while a truly reliable mean might require more iterations than are feasible, my hypothesis is that several rounds (i.e. approximately 8-10) would generate enough stability to make the process worthwhile, especially given other effects brought about by iteration (e.g. trust-building).

In a sense this approach to iteration in deliberation localizes to the individual and to particular groups of participants in a deliberation the stabilizing effect based on an argument offered by Page and Shapiro (1992): namely that, in general, large groups of individuals

approximate accurate, stable opinions by washing out ill-informed, uninformed, or random positions. One challenge to this approach is that some biases are *structural* so that large numbers of participants or iterations by a single participant cannot be washed out simply by group size or repetition since the stable nature of the bias supersedes randomness (Kuklinski and Quirk in Lupia et al. 2000). Moreover, given what we know about mimicry and affective contagion—in essence, the matching of behaviour of spreading of an anchoring feeling—it is plausible that individual instances of random change can, especially in a close-quarter deliberative exchange, be transmitted to others, as we saw in the example above. This risk is what makes the challenge a structural one: strong mechanisms will emerge that can bias proceedings in a particular direction that subsequently spreads to others (individuals and groups).

Iteration is a tactic employed to create central tendencies over time that will help to reveal the often-random nature of certain opinions and preferences *before* a final judgment or decision is made. This may occur *after* cascading, structural biases emerge, but can help keep these preferences from becoming more-or-less fixed through exposure to competing information. However, this tool is not, on its own, enough to address the threat of randomness to the generation of epistemically good decisions. What if an individual's bias generated by a-rational cognition and cognitive distortion *before they even begin a deliberation*? If that is the case, then the threat of a spread of bad information precedes the deliberation. What if the speaker is particularly eloquent or the misinformation is particularly pervasive or persuasive, or if the biases brought into the room are *themselves* structural rather than random (which is

likely to be true in the case of motivated social cognition or moral intuition)? The third tactic for maximizing institutional deliberative autonomy is designed to address this.

### *A-rational receptivity*

One of the significant challenges posed by a-rational cognition and cognitive distortion is non-conscious nature of how its processes operate and effect behaviour. By definition, these challenges exist liminally, just outside of awareness, but not below the threshold of having an effect. Training oneself to be aware, as much as possible, of what contributes to your judgment is difficult. Moreover, while, as we have seen, it may be impossible to be aware of *each and every* factor that goes into producing a judgment, *especially* those that operate at non-conscious level, it is much easier to ascertain one's current mood and the emotions associated with it at any given moment.

In light of this, rather than attempting to bury a-rational cognition and the threat of cognitive distortion, rather than pretending they do not exist, or attempting to approximate the rationalist ideal of "the econ," deliberative citizens should try to engage with a-rational cognition and states in an attempt to, as best as possible, bring to the surface hidden, partial, or manifest moods, feelings, emotions, motivations, fears, and anxieties that rest at, near, or even deeply below the surface. Towards this end, a-rational receptivity refers to a group-wide disposition aimed at encouraging the exploration of an individual's affective state. Such an element is institutional because it is built into the structure of the deliberation and stipulated as necessary; however, its operating mechanisms are individual practices of self-interrogation combined with group-wide receptivity to such interrogations.

Does such deliberation, then, become a kind of group therapy—a kind of collective exploration of how affect influences our judgments and decisions? Ever so slightly, perhaps, but with a specific end in mind: the examination and elaboration of individual motives to the best of the deliberative citizen’s ability. Moreover, even if this approach borders on the therapeutic—and who says it should not?—it stands to reason that it is nonetheless more desirable to understand and make public, to the extent that it is possible, the factors that go into motivating an individual to come to a particular judgment (whatever they may be). In so doing, the ideal of autonomy is better approached, since as more motivations and reasons become manifest, they can be interrogated, challenged, revised, affirmed, or dropped all together, which is essential to good democratic deliberation. After all, suppressing such affective content, imagining it does not exist or trying to eliminate it all together is useless at best and likely quite deeply counterproductive.

As Damasio (1994, 2003) has shown, feelings and emotions (what I lump together, along with moods, as “affect”)<sup>54</sup> do important cognitive work; and as Heath (2014) and Pinker (1997) have neatly summarized, there are good evolutionary reasons why affect looms large in our lives. When it comes to the force of affect, we ought to be able to row with the current when necessary, though we should also work hard to know where it is taking us. Returning to Kant: autonomy requires that individuals have both choices and reasons for choices; I hasten to add that the practice of autonomy for deliberation requires that individuals have *valid and accurate*

---

<sup>54</sup> I am very much aware of the large body of literature known as ‘affect theory’, though I have chosen to adopt the usage and approach from mainstream cognitive science, as I have noted. This is because I am interested in affect—feeling and emotion—specifically as they relate to cognition in an epistemological sense. The literature with which I engage tends to use affect in this sense, and so I have adopted the same convention.

reasons for choices, which would include the emotions, feelings, and moods that play a role in generating and perhaps sustaining those choices. The question then becomes “How can we do this?”

At its most basic level, a-rational receptivity requires individuals to be open to justifications for preferences that have affective content. For instance, an individual who registers "disgust" as their reason or justification for opposing some policy (say, for example, polygamy) is *prima facie* violating the principles of public deliberation since one's disgust is not a public, accessible reason. Nonetheless, as many researchers have demonstrated (see for instance Damasio 1994, Haidt 2003, and Schnall et al. 2008), there is coherent *content* in the utterance that one opposes some practice, like polygamy, out of sheer disgust. Being *open to* receiving a-rational (i.e. affective) reasoning does not require *acceptance*, but it does require an openness to receiving such a "reason"—and then, for the purposes of deliberation, pushing back and searching for deeper reasons. If none exist—or none can be ascertained—those who deliberate can bracket the issue or the intervention and move on. At the very least, an explicit openness to receiving a-rational justifications help prevent individuals from smuggling in rationalizations in an attempt to avoid violating the deliberative norm of rational, public reason giving.

In short, being open to a-rational expressions encourages individuals to be more honest about their reasoning (at least to the extent that such a practice is possible). Merely having, for instance, a declaration of disgust on the table makes it scrutable and thus an item for further discussion and deliberation; leaving it hidden closes off avenues of pursuit that might be useful for deliberation down the line. Thus, individuals can be a-rationally receptive, at a basic level,

by being (at least provisionally) open to reasons/reasoning that are primarily affective. (Of course, individuals should also be attentive to the possibility that *no reasons* exist for one's support or opposition to a practice or policy—or, at least, that none exist that are likely to be accessible.) Admittedly, this requires a certain openness to creativity, but individuals are often what psychologists call “creative thinkers” (Torrance 1959); as Pizarro et al. note

...mounting evidence suggests that an exclusively reason-based view of moral judgment is wrong a psychological theory. Not because people do no reason at all when they make moral judgments...but because other processes are at work as well.<sup>55</sup> There is evidence that everyday moral judgment is a much less rigid, more emotional, and more flexible process than previously described... (81)

Accordingly, if what we seek from deliberation is, at least, open and honest exchanges, then those who deliberate—as well as those who design deliberations and those responsible for developing the normative expectations which deliberators are expected to strive for—ought to be open to a-rational thinking and reasoning.

### *Cognitive diversity*

One significant challenge towards autonomy in deliberative contexts emerges from the potential effect that a-rational cognition has on an individual's creativity: coasting through deliberations by relying on automatic, low-level or non-conscious heuristics or gut-level

---

<sup>55</sup> For instance, expressions of empathy.

guidance may often be useful for navigating a complex, high-speed, and routinely variable world, but the kind of attention and reflection that deliberation requires is usually poorly served by this approach. Cognitive diversity<sup>56</sup> may hold part of the answer to getting subjects out of this state. Earlier I defined cognitive diversity via Landemore (2013), who draws on Hong and Page's (2004) conception, and suggests that it "refers to a diversity of ways of seeing the world, interpreting problems in it, and working out solutions to these problems. It denotes more specifically a diversity of perspectives...interpretations...heuristics...and predictive models" (see also Page 2007).

In deliberative contexts, the presence of a variety of individuals might not only improve decision making (Page 2007) and the quality of deliberation (Landemore 2013), but that presence of diverse ways of thinking may also offer a cognitive jolt to individuals who might otherwise rely heavily on the low-resource flow of automaticity when processing information and coming to judgments. In order to disrupt a-rational coasting—cognitive autopilot—it seems necessary to remove individuals from the comforts of familiarity and habit: a goal that might be well-achieved by the presence of alternative approaches to examining information and arguments, since participants will be less likely to piggy-back on the thinking and arguments of others (which, in theory, should also boost the claim to be acting autonomously, rather than simply following a leader or set of leaders).

---

<sup>56</sup> In theory, cognitive diversity does not necessarily require a diversity of identities among those who deliberate, only a diversity of approaches towards cognition and its related practices. That said, to the extent that differences in identity correlate with cognitive diversity, the two can be thought of as bound up together in a sense.

Research from management theory suggests that cognitive diversity can be an asset in decision-making contexts in which “task conflict” might arise—that is “disagreements about the issues such as appropriate choices of alternative policies or differences of judgment about the decision” (De Dreu 2006, cited in Olson et al. 2007). Cognitive diversity has also been found to correlate positively with mediating task conflict (Olson et al. 2007). This study, while limited in scope to management executives, nonetheless offers proof-of-concept for the ability of cognitive diversity to prepare a group to make a complex decision. By implication, disagreement brought about by cognitive diversity challenges the cognitive auto-pilot that threatens to undermine autonomy. This is especially true in cases where there are high levels of competence-based trust—a good that might be brought about in the context of democratic deliberation through iteration (Olson et al. 2007).

As we will see below, getting individuals to a point where they are motivated to pay attention and engage requires that one or more of several motivating factors be present. Cognitive diversity might well be one of those factors. At this stage, however, when it comes to democratic deliberation the cognitive diversity hypothesis is mostly still just that—a hypothesis—and more (careful) research is required. Nonetheless, if having diverse ways of approaching a problem means that individuals are more inclined to be critically engaged, and to consider closely their reasoning pattern, then the idea seems promising; however, researchers and practitioners should remain cautious about the degree to which cognitive diversity is presented as a tool for enhancing democratic deliberation. On balance, studies into the effects of cognitive diversity have yielded mixed results, and a coherent definition of the concept (and

a conception that lends itself to operationalization) is required (see, for instance, Mello and Rentsch 2015).

### *Targeted motivation*

As we saw in chapter three, in the 1980s, psychologists Richard Petty and John Cacioppo developed the elaboration likelihood model (ELM), which uses dual-process theory to specify two general routes through which a statement or argument might be processed: the central route, along which subjects were more likely to scrutinize a message, and a peripheral route along which subjects were more likely to employ cognitive short-cuts and external cues to evaluate it (Petty 1999; Petty and Cacioppo 1981, 1986). As the authors discovered, the key to getting subjects to employ the first route—one far better suited to the goals and exigencies deliberative democracy—is motivation: various factors, including a message’s relevance or the availability of cognitive resources, went into determining which route a subject was likely to take (Chaiken and Trope 1999; Jae and Delvecchio, 2004; Petty, Wells, and Brock 1976).

Targeted motivation is an attempt to engage central-processing by highlighting to participants in a deliberation the relevance and importance of an issue. When it comes to a-rational cognition and cognitive distortion and their impact on autonomy within the context of deliberation, targeted motivation may assist in shifting an individual’s attention towards the subject matter and *towards their way of thinking about it*, thus maximizing the likelihood that they will scrutinize the data and arguments presented to them. Such scrutiny, in theory, would also open individuals up to iteration and a-rational receptivity, as well as traditional reasoning, by raising the cognitive stakes of the deliberation and reminding them why their engagement

matters—not just to the broader community, but to themselves. Ultimately, targeted motivation alone may not work on all of the deeply embedded processes of automaticity, motivated cognition, system justification, and social intuitionism (i.e. processes that impact deeply at both first- and second-order levels); it may, however, minimize the impact of some of the superficial effects of automaticity and may enable deeper tactics, such as a-rational receptivity, to work. And recall: cognitive autonomy, especially as it relates to institutional deliberative autonomy, is not evaluated as a binary state of on/off, but rather as falling somewhere along a continuum ranging from less autonomous to more.

Research on the ELM and brand attitudes/intentions is revealing. For instance, Mackenzie and Spreng (1992) find that by increasing motivation to evaluate (in this case, subjects were encouraged to review and evaluate an advertisement), central processing was strengthened, and peripheral processing was weakened, interestingly, “by influencing the strengths of (central) relations...rather than by influencing the amount of central processing or mean peripheral cue perceptions...” (527). Thus, targeted motivation to evaluate something (in this case, an advertisement for a watch) may not lead to *more* central processing (although it might); however, it does increase *the effect* of central processing while weakening peripheral processing. In the case of democratic deliberation, this is encouraging—a combination of practices (targeted motivation and others) may likely work together to enhance central processing and thus support autonomy in such a way as to enable individuals to produce better judgments and groups to make better decisions.

Targeted motivation can be further subdivided into tactics aimed at engaging individuals. Specifically, targeted motivation should take the form of ensuring that: i) arguments

are presented in clear, manageable form and language; ii) individuals are given appropriate amounts of time to scrutinize information, ask questions, and discuss their perspectives; iii) rewards for adopting peripheral methods are minimized or eliminated (e.g. rewards for finishing early or before another group or sub-group); iv) individuals are presented with clear arguments as to why a given issue is relevant to them, their families and friends, their community, city, state, or country; and v) the environment in which deliberation occurs is free from distracting elements, including any irrelevant stimuli that may provide subtle nudges as to which way a participant should decide.

### **Next steps for deliberative institutional design: hypotheses and tests**

As noted, institutional deliberative autonomy is at once an aggregate-level evaluation tool for autonomous deliberation and a series of concepts designed to help generate more autonomous deliberation from and among individuals. In sum, it is a standpoint from which one can judge democratic deliberation while also serving as a program for encouraging more rational, autonomous deliberation. The concepts I have offered here that inform that standpoint are mostly just that—concepts. They have been drawn from the work of other political theorists and political and social psychologists, and have been theoretically applied to the challenge and hope of autonomous political deliberation aimed at generating epistemically good judgments and outcomes. The next step for IDA and the concepts attendant to it is the development of specifically related hypotheses that can be tested in real-life deliberations. This requires that these concepts be used to generate specific institutional design plans and that these plans, or

elements of them, are attached to hypotheses that can be evaluated as producing more or less autonomous judgments and outcomes.

These specific hypotheses, once formulated and tested, will yield empirical data that can be used to generate further hypotheses and institutional design concepts and plans—and would also help rule out those that do not work well or at all. Attention to design will generate a cycle/feedback loop of concept-hypothesis-test-design that will be beneficial to those who deliberate, those who design and facilitate deliberation, and the public at large, who will benefit from the generation of better judgments and decisions. This commitment, however, is the work of another project, and perhaps alongside another researcher or group of researchers. Determining which specific hypotheses should emerge from these concepts, deciding how specifically they should be designed to fit into a deliberation, and choosing how to evaluate them will be a challenging, but important project. This chapter has made a case for why institutional design is well-suited to responding to some of the challenges of a-rational cognition and cognitive distortion to epistemically good judgments and decisions, outlined why IDA and its related concepts are important to better deliberation, and has presented four concepts that can be used to generate testable hypotheses.

## **Conclusion and Summary**

In this chapter, I have argued that to address the challenges that structural cognitive distortion bring to deliberative democracy as a theory of epistemically good decision making, those who design deliberations must use institutional designs and personal practices to begin to overcome certain cognitive limitations. I have also introduced the concept of institutional deliberative

autonomy (IDA) as a way to think about autonomy in deliberations. I have suggested potential measurement/evaluation tools for ascertaining when and to what extent autonomy is present during deliberation—the Discourse Quality Index, expert qualitative reports, and self-reporting. Furthermore, I have examined four concepts useful for generating the aforementioned institutional designs and personal practices—iteration, a-rational receptivity, cognitive diversity, and targeted motivation. Finally, I have outlined the steps that theorists of deliberation and practitioners should take to begin to test approaches towards generating better deliberation.

## Chapter 6: Democratic deliberation and institutions

One of the questions I ask in this dissertation is: What is the impact of institutions on cognition? While I am most interested in asking whether citizens can make rational, autonomous, and thus epistemically good, judgments and decisions within deliberative contexts, that requires an assessment and exploration of the challenges that emerge from both *within and outside the individual* in the pursuit of such judgments and decisions. Challenges from within have to do with both how internal cognitive states affect judgments and how internal cognitive processes are affected—and compromised or manipulated—by external phenomena. Challenges from outside arise from the ways that institutions have been set up or have evolved and how they impact cognition—including how individuals or groups might leverage certain institutional features to manipulate outcomes deliberately. In essence, given these phenomena, the specific question of note that emerges is: How and when do institutions encourage us to think, judge, and decide poorly? In this chapter, I will move away from individual cognition and the subjects of rationality, autonomy, and epistemically-good judgments and decisions, and I will focus on the broader social, cultural, and political context that informs this question: the institutional frameworks of our political life.

Specifically, first, I define what an institution is, and give some examples of the institutions that undergird our world. Second, I discuss how institutions and cognition interact. Third, I outline the ways that institutions encourage poor thinking—that is, the ways institutions, through deliberate manipulation or incidentally, distort cognition and lead to sub-optimal judgments and decisions. Finally, I put the question of the effects of institutions on our cognitive in the context of deliberation and explain how theorists of deliberation might respond

to the challenges mentioned. Then, in chapter seven, I discuss how deliberative systems theory might respond to the challenges raised in this chapter.

## **What is an institution and why do institutions matter?**

In the following sections, I explore overlapping and intersecting conceptions of what precisely comprises an institution. Before getting into the various specific aspects of what makes an institution, however, I want to provide a brief overview of the concept of an institution and where it fits into this project. The very idea of an institution is an amalgam of rules, norms, expectations, and practices, as well as resources and power—not to mention rewards (for respecting norms) and sanctions (for violating norms). Institutions collect and organize practices and in doing so help pattern behaviour—or help formally codify or informally outline behaviour. Because the concept of an institution is fuzzy, however, institutions risk becoming everything and anything. In the sections below, I try to specify what I mean by an institution to avoid this problem.

Institutions fit into this project because not only do they affect the context in which political actors make decisions, they can also help correct for mental limitations (as I argue in depth below). As some researchers have argued (e.g. Heath 2014), our institutions are extensions of ourselves that go beyond—that is, enhance—our natural abilities. In this way, *cultural evolution* has overtaken *biological evolution* as the driver of our progress as species. For my purposes in this chapter, I will focus on political institutions—how they assist our judgment and decision-making and how, in some cases, they detract from it.

### *The rules of the game: Denzau and North*

Denzau and North (1994) define an institution as “...the rules of the game of a society [consisting] of formal and informal constraints constructed to order interpersonal relationships” (4). They contrast institutions, which are external to the mind and which assist humans in collectively ordering the world and cooperating with one another, with mental models, which are “internal representations that individual cognitive systems create to interpret the environment” (4). Mental models are “mind created,” while institutions are “individual created”—that is, they are generated through the actions of individuals. Together, along with ideology, mental models and institutions are how humans structure and make sense of the world. Knight (1995) argues that institutions serve us by decreasing uncertainty, providing information, and structuring formal and informal sanctions for those who participate in them. Turner (1997) defines institutions, more specifically “social institutions,” as “...a complex of positions, roles, norms, and values lodged in particular types of social structures and organizing relatively stable patterns of human activity with respect to fundamental problems in producing life-sustaining resources, in reproducing individuals, and in sustaining viable societal structures within a given environment” (6).

Summarizing and relating the above definitions, we can say, for our purposes, that institutions are structures within which human interactions occur, which are generated and enabled by repeated, patterned human activity, and within which mental models are mobilized and presented for collective consumption and interrogation towards some end, and which organize, enable, and direct human interactions through creating incentives and disincentives, expectations, rules, frames, and power disbursements. It is also important to note that

institutions are reflexive—they shape behaviour, but they are also interpreted, understood, and internalized by individuals who make sense of them and thus give them form, ensuring that they continue as long as there is a critical mass to enable them to order and make sense of the world.

So, institutions are essential for life, but they are not neutral. Moreover, they are subject to hijacking and can serve to bring about sub-optimal judgments and decisions. As Denzau and North (1994) note, there are challenges to substantive rationality—and, I would add, to autonomy—embedded in human interactions, including institutions, that revolve around three categories of condition: complexity, motivation, and information quality/frequency. That means that while institutions play an essential role in organizing life, they can be structured in ways that either enhance or detract from the kinds of behaviour that may be desired by interacting with individual human capacities and cognitive tendencies (as we will see below).

Looking at institutions and how they interact with cognition, autonomy, and rationality in *individual choice scenarios*, Denzau and North first look at *complexity*, asking: “How complex are the mental models required to make sensible choices given one’s preferences and resources?” (Denzau and North 1994: 7). They determine that within institutional contexts, individuals tend to swap in familiar mental models when presented with novel ones, relying on what they already know to make new decisions. Next, they look at *motivation*, in which two aspects are considered: the importance of the decision to the individual and belief that one’s individual choice matters regarding affecting the actual outcome of the issue at hand (7-8). Finally, *information quality and frequency* refers to whether or not the data available to the

individual is enough and good enough for them to make judgments that are corrective of bad models—as a feedback provision (8).

### *The ontology of institutions: Searle*

The philosopher John Searle has also sought to answer the question ‘What is an institution?’. Searle approaches the subject conceptually. Arguing that the literature from Aristotle to the present has done a poor job at clarifying *precisely* what the concept entails, he puts the question like this: “What is the ontology, mode of existence, of institutional reality?” (Searle 2005: 1). His inquiry stems, largely, from questions of economics and the social and political functions of aggregative economic behaviour, but his emergent understanding of institutions are, by and large, transferable. While an exposition of Searle’s extended definition of an institution, which is precise and carefully constructed, is outside of the scope of this chapter, and indeed outside the purposes of this dissertation, one element of his understanding of institutions is helpful for understanding how variable institutional setups are important determinants of behaviour.

Searle argues that the logical structure of institutions is “X counts as Y in C.” His “special theory of the logical structure of institutions” suggests that: Some feature assigned to some person, object, or phenomenon (X) is granted a special status (Y), in a given context (C). All of this exists *outside of* the physical properties of the thing itself. So, for instance, money is granted a special status of having the power to be used to purchase a good or service, and this has nothing to do with its physical form (especially since you can have money without currency, such as with a debit card) (Searle 2005: 5-10). Searle makes what he calls a “strong claim,”

arguing “the institutional ontology of human civilization...is a matter of status functions imposed according to constitutive rules and procedures” (Searle 2005: 9). This institutional ontology then creates power relationships that enable human cooperation in ways that are impossible without these deontic (i.e. rights, duties, obligation generating) powers (Searle 2005: 10). Once institutions are established, Searle concludes, they then further enable institutional facts to emerge within their structure (e.g. most of the time, crossing the goal line with the ball in football counts as a touchdown).

What is interesting about this argument and the conception of institutions that underwrite it is that Searle’s focus on the *assignment* of status functions reveals that institutions are *human made* and that their function is to enable certain realities to emerge and to encourage what he calls “desire-independent” reasons for action (i.e. reasons for action that exist outside the moment-to-moment desires or inclinations of individuals). Part of the challenge of institutions to cognition is that the development of particular institutions and the assignment of enabling status functions may or may not encourage rational and autonomous cognition. Institutions enable certain behaviours that outpace what would be possible without them, but they also create heuristic structures based on trust relationships or expectations that encourage status quo thinking and behaviour. Or to put it simply: institutions tend to be conservative by definition. All of this is fine of course, except that once the enabling logic of institutions is set, the social realities that emerge around those institutions take on a life of their own (for better or for worse), especially as those who benefit from them work to preserve and entrench them. Of course, by allowing humans to coordinate action over time and space, institutions serve an essential role in creating stability and continuity—hence the need for them

to typically be conservative. However, despite this good, as I will argue further below, when it comes to generating better judgments and decisions through deliberation, some existing institutions may occasionally work against the project through entrenching incentives and disincentives that, while serving certain interests, may not encourage autonomous, rational judgment, even if they do important work in ensuring stability and continuity.

Perhaps what is required to begin to address the challenge, then, is an alternative approach to conceiving of institutions and how they relate to our behaviour within them. Max Cameron (forthcoming manuscript) extends our understanding of institutions, adding that while they generate rules and incentives, thus enabling and constraining behaviour, they also generate *practices* that create *social goods*. As Cameron puts it: “Institutions institutionalize practices” (6). Cameron undertakes this extended understanding of institutions within the context of his argument that narrowly conceiving of institutions in a rational choice framework focuses too much on rules and incentives under the rubric of competitive utility maximization (12). Such a narrow focus undermines an understanding of institutions that includes their (potential) function as generators of skills and enablers of capacities—in general, understood as ‘practical wisdom’ or the ability to ‘know what to do, when to do it, and how to do it’. Or, to both summarize Cameron’s point and relate it back to the issue at hand, institutions play an essential—if sometimes underappreciated or underdeveloped—role in enabling individuals to make good judgments and decisions through the cultivation of “cognitive and ethical skills” (14-15). This institutional function is a particularly important one since, as Cameron argues, rules cannot anticipate all likelihoods and contingencies, which is why good judgment is important (15).

If we imagine rules as (among their primary constitutive or regulative functions) strong heuristics—after all, they are, among other things, ways to ‘solve problems’—and if we also accept Cameron’s argument that our focus on institutions should move beyond conceiving them merely in terms of rules (and sanctions for rule violations) and towards an understanding that includes the skill and capacity generating elements of them, then institutions may become important correctives to cognitive constraints. How? Well, if part of what drives cognitive distortion is a tendency to rely on poor thinking habits, including heuristics susceptible to manipulation, and if institutions can act as (moderate) correctives to this by generating more refined thinking skills, then this new conception of institutions may be essential for addressing the threats to autonomy and rationality broadly conceived. To return to Cameron, he argues that institutions can be set up in such a way that enable individuals to cultivate good judgment and related practices.

This argument leads us to the question of how we might set up institutions so that they encourage better thinking—that is, thinking that supports and encourages autonomy and rationality—and, subsequently, better judgments and decisions. To revisit a key issue raised earlier, in chapter one: the conceptions of autonomy and rationality that I am drawing on *are not derived from the rational choice tradition*. I am not just speaking of maximal autonomy or ends-means rationality or else rationality understood as a capacity for maximizing utility—though either of these approaches to conceiving of human rationality may, at times, be the best way, in a given context, to conceive of rationality as a guide or goal. But in the context of real-world political judgments, I think of rationality as far more than the capacity to reliably and consistently gather information about the world and to match that information to intentions—

whatever they may be. Rationality has to do with not only a capacity to collect and understand information from and about the world; it also requires that individuals possess a capacity to develop reasons for their preferences and actions, to support those reasons through sound and valid reasoning by interacting with autonomy, and to thus produce judgments consistent with these capacities. Indeed, these capacities are what help make institutions possible—as well as legitimate and coherent.

I think of autonomy as the capacity to produce, scrutinize, and affirm choices, goals, and values in such a way that reflects an awareness of the factors that go into informing the reasons, reasoning, and judgments of an agent or the decisions of a group. Thus, autonomy enables rationality. To put it very simply: autonomy and rationality, in the sense that I am interested in them in most real-world political and social contexts, are about asking whether individuals can make consistent and reliable sense of the world and know what is driving the intentions they have that come from this sense-making. But before getting to the question of how to better set up institutions to support good judgment and decision making, it is important to understand first how institutions and cognition interact, and, second, the particular ways that institutions, *especially expressly political institutions*, encourage poor thinking.

### **How do institutions and cognition interact?**

Returning to a point raised in chapter five, the principle that *structure may dictate function* is critical to understanding why institutions matter to the kinds of judgments and decisions we make. Institutions do not just set out idle rules and incentives; at both the cognitive and behavioural level they shape our behavioural dispositions and set the cognitive frames and

narratives we use to sort, understand, and navigate the world—pushing and pulling our thinking, enabling and constraining certain ways of seeing the world, rewarding us for certain things and punishing for others. I would go so far as to say that institutions, broadly conceived, and including social, political, and economic forms, *make up our primary way of being in the world*. When it comes to deliberative democracy, if this is true at the level of specific deliberative set-ups—the many ways of “doing deliberation”—then it is just as true as these smaller set-ups are scaled up, arranged into systems, and as they operate alongside other large institutions. In this section I will focus on how institutions—again, understood broadly—interact with cognition: How they shape and distort it, how they enable certain ways of thinking and constrain others. To do this, I will look at two psychological concepts linked to institutions: ontological security and meaning maintenance, each of which relates specifically to institutions and *cognition* at the point *before* any particular act or judgment.

Before moving on, though, I want to say more about this last point. Cognition is not an institution itself, at least not in the sense that I mean institution here. Cognition is a phenomenon or process that enables institutions; through facilitating human activity, it is in turn affected by the institutions enabled by it. The way we process information, the kinds of information we process, the sorts reasons we give, the types of reasoning we engage in, and the judgments we make will be affected by the institutions that exist in the world, giving the environment in which we live shape and making it more or less predictable. The elements that go into generating this shape and predictability include the incentives, disincentives, rules, frames, and narratives that are generated by and embedded within these institutions. Again, recall that structure may dictate function: we can apply this principle at both the level of

behaviour (i.e. conceived of as action) and cognition (i.e. conceived of as how we collect, process, and sort information as we generate behaviour). Institutions, insofar as they interact with cognition and shape behaviour, are profoundly powerful. The more we understand this power, the better, since better understandings will open up ways of thinking about shaping institutions so that both better ways of behaving and thinking can be generated.

*Ontological security: What is there in the world?*

Part of what makes institutions essential to human life is the *psychological* effect they have on those who benefit from their existence. While institutions have what we might call “tangible” effects—that is, they produce obvious, material, day-to-day outcomes, such as facilitating exchange or travel, or peacefully transferring power from one political party or ruler to another, or allowing a sporting match to carry on efficiently and coherently in a large group—they also have a subtle effect on humans. When it comes to psychological goods produced by institutions, first and foremost, they allow for the establishment and maintenance of *ontological security*. Giddens (1984) defines ontological security in relation to an emergent phenomenon linked to institutions. If institutions are rules that enable social structures, then ontological security is the emergent cognitive state that recognizes those social structures as relatively fixed and stable. That is to say that ontological security refers to a state of mind (or a sort of psychological homeostasis) that reflects the confidence we have that our day-to-day lives will remain more-or-less the same and that the social world will remain coherent and recognizable across space and time. This, of course, is *often* one of the desirable effects of institutions: they enable ontological security. But it is not *always* the case that institutions are

socially desirable, fair, or just. Slavery is one of the most odious institutions that human beings have come up with, and one of the world's and history's most secure; the practice of slavery pre-dates recorded history and continues today with as many as 36 million people currently enslaved.<sup>57</sup> So while institutions generate reliable and familiar contexts, environments, incentives/disincentives, ways of thinking, and behavior, they are not universally good or desirable. None the less, they are essential parts of the scaffolding of our day-to-day reality.

Institutions enable ontological security by establishing reliable patterns of behaviour and outcomes. However, the stabilizing function of institutions as they underwrite ontological security, it seems to me, allows for exploitation, since they are open to being manipulated in such a way that human judgments become distorted (e.g. by normalizing and justifying a practice such as keeping slaves). This is especially true in relation to the human psychological tendency to seek a narrative explanation of phenomena (Kahneman 2011). In the end, it does not seem to be the case that humans are fundamentally *truth-seeking* beings as much as we are *meaning-seeking*—although this meaning might be made into truth over time through embedding meaning in our concepts, norms, and institutions. As Nietzsche (in Kaufmann 1977) writes about truth, it is

A mobile army of metaphors, metonyms, and anthropomorphisms—in short, a sum of human relations which have been enhanced, transposed, and embellished poetically and rhetorically, and which after long use seem firm,

---

<sup>57</sup> This number, from late-2014, includes those “subject to forced labour, debt bondage, trafficking, sexual exploitation for money and forced or servile marriage.” See: <http://www.bbc.com/news/world-30080578>.

canonical, and obligatory to a people: truths are illusions about which one has forgotten that this is what they are; metaphors which are worn out and without sensuous power; coins which have lost their pictures and now matter only as metal, no longer as coins (46-47).

Regardless of the nature of *truth*, the fact that human beings regularly (and sometimes desperately) seek *meaning* and *stability* is what should push defenders of deliberative democracy, especially those who defend it on epistemic grounds, to seek *better systems* rather than just better individual practices or deliberative design. If deliberative democracy is to provide a major alternative—or even just complementary practices—to aggregative democracy and other approaches to public decision making, then theorists must take into account the way democratic deliberation *as an institution* will respond to and exist with other institutions at the system-level (a subject I address in chapter seven). While it seems unlikely that any combinations of institutions will ever be able to eliminate distorted judgments, it is obvious that they can do better or worse as facilitators of good judgments and decisions.

### *Meaning maintenance: Start making sense?*

Another related and essential way that institutions interact with human cognition is through mediating our structures of meaning. If ontological security is about the stability of what exists in the world, meaning maintenance is about *how we perceive and understand the content of what exist*. More specifically, meaning maintenance is concerned with “...the manner in which

people respond to violations of their committed meaning frameworks [and how they] restore a sense of familiarity to their experiences” (Proulx and Inzlicht: 2012: 317).

Proulx and Inzlicht define meaning as “the mental representations that allow us to understand our experiences, whatever they may be, and however, they may come to be understood” (318). They note “At its core, meaning is the expected relationships that allow us to make sense of our experiences” (2012: 317). Violations of these expectations generate somatic and psychological responses (such as anxiety). The authors further suggest, building on the meaning maintenance model developed by Heine, Proulx, and Vohs (2006), when violations of meaning structures occur or in times of uncertainty, individuals attempt to reconcile those violations through “fluid compensation”—the swapping of one meaning framework for another while attempting to secure a sense of meaning in a given context as we struggle to determine *why something happened* and *how it will affect us and the world*. Meaning is thus about, as Proulx and Inzlicht note, “sense making” (319). We can see how meaning and sense, as related here, connect back to institutions, especially to Searle’s deontic conception of an institution as including expectations embedded in rules. Rules enable stability and predictability, and insofar as they become agreements (tacit or explicit) on their own or set into larger institutional structures, they enable sense making.

Take, again, the example of currency, which is a sort of institution. While individuals exchanging pieces of coloured paper for goods and services is in and of itself meaningless, the agreement that *this particular sort* of coloured paper, when printed by a federal body with permission to do so, and circulated by a central bank, is currency that is part of an economy makes perfect sense and could easily be apprehended (or explained) to someone unfamiliar

with the system. However, if, tomorrow, for some reason, a group of people started exchanging different pieces of paper randomly and for no discernable reason or return, we might stop making sense of things. This would be because we would have no reference to an enabling and sense-providing institution.

We can also imagine our meaning structures as related to ontological security as the cognitive representations of what we come to think of as existing in the world and our understanding of how those things fit together. Moreover, we can imagine both ontological security and meaning maintenance as serving as the enabling factors for institutions. After all, institutions are largely sets of rules, guidelines, practices that give meaning to phenomena, generate patterned outcomes, and guide behaviour by enabling and constraining possibilities in a given context. I see institutions as the emergent, real-world, rule-based or practice-based manifestation of our drives towards ontological security and meaning maintenance. But they also *embed power relations*, such that the rules, practices, and meaning/sense that come from institutions also shape and reflect distributions and practices of power, which I have argued, also shape cognition by privileging some ways of thinking over others, enabling certain frames more than others, and, when it comes to political practices such as deliberation, by playing a role in determining who gets to the table in the first place and who gets to speak. So, this means that institutions are not merely *ways of doing things*, but rather are essential, life-enabling elements of our reality that are conditioned by, and that condition, relations of power. While I am not primarily concerned here about the flow of power within institutions, aside from the particular effects of certain practices or phenomena on cognition, I want to highlight the power element of institutions and avoid psychological reductionism when it comes to

understanding them. After all, power is *enabled* by institutions and by asymmetries in *resources and capacities* (e.g. money, education, personal networks, quality of writing and speech) embed power over how institutions function.

Now, however, essential they may be, certain institutions—perhaps most or even all—are subject to shortcomings, manipulation, distortion, hijacking, and outright failure. In this dissertation in general, and in this chapter specifically, I am most interested in the way that institutions come up short in the service of generating good outcomes, despite the good they do, and either generate or enable or contribute to sub-optimal individual judgments and collective decisions. In the following pages of this chapter, I will discuss the particular ways that certain institutions can generate poor outcomes in a *political* context. So, before moving on, in chapter seven, to a discussion of how deliberative institutions, as part of a deliberative system, might encourage better judgments and decisions, I will first explore how institutions affect judgment in a negative way—that is, a way that minimizes or reduces opportunities for autonomous, rational judgments and decisions. While institutions also clearly serve an essential positive function—along with other practices and systems—in enabling the extension of human consciousness into the world (Heath 2014), it is their shortcomings that interest me in the context of this project. Again, it is important to understand that the argument I am making is not that institutions *are universally bad at encouraging good judgment and decision making*, but rather that *we can do better at leveraging institutions (through design) to encourage better judgments and decisions*.

## **How do institutions encourage poor reasoning, judgments, and decisions?**

In this section, I will examine five common and significant *political* institutions and explore how they interact with cognition, autonomy and rationality to affect judgments and decisions. These institutions are common insofar as, whatever their particular manifestations globally, they are found in one form or another in contemporary liberal democracies. They are significant because each plays some important role in shaping or determining public outcomes by affecting individual judgments. The examples I use are drawn from the Canadian and American political experiences, and each is intended as evidence in building the case for a proof of concept: that is, as evidence that demonstrates how political institutions *can and do* encourage, in some instances, poor judgments by exploiting—through deliberate manipulation or otherwise—certain common cognitive shortcomings. Not all the institutions that I discuss do this to the same extent or in the same way; however, it remains important to understand how and why those that do can do so to improve both personal practices surrounding them and the institutions themselves. This section will move away from deliberative democracy and to more common and traditional democratic institutions. I will then return to deliberation in chapter seven—specifically to deliberative systems—to ask how deliberation as an institution might complement and help improve these other kinds of political institutions.

Political institutions are important to look at because they help shape not only our social and political lives but also how we deliberate within them. Deliberation does not occur in a vacuum; rather, it is a way of doing politics that finds itself alongside elections, parliaments, protests, and other forms of political activity. So, understanding how these political institutions shape thinking and behaviour will help us understand how we might conceive of deliberative

democracy and how we might design deliberation to maximize the generation of good judgments and decisions. It is to these complementary political institutions that I now turn.

### *Campaigns and elections*

The sheer volume of literature on political campaigns is overwhelming: how and why people vote or do not vote has been a subject of systematic study for decades. The study of campaigns and elections include the heuristics individuals use to make decisions, how political information is process, the reliability and stability of the electorate as individuals and as a whole, and more. My focus in this section is specifically on how campaigns and elections—not electoral systems themselves—as sets of rules about how candidates interact with voters and how the media treats each affect the way individuals think. I will draw on literature from Canadian and American political psychology to show that the ways of thinking encouraged, in general, by these set-ups tend to generate decisions and outcomes that are sub-optimal. This point is important and applies to each of the institutions that I will look at in this chapter since it highlights the fact that my argument is that *it is the institutional setup that encourages poor thinking*. It is not that individuals are incapable of making good decisions—that they are “too dumb”—but rather that the institutions that surround them and with which they interact cognitively bias judgments and decision making in poor directions by exploiting our cognitive limitations. And all of this is made to seem much worse given the gap, discussed a few times in the first half of this dissertation, between our normative expectations for ourselves as individuals and as democratic societies, and our cognitive capabilities.

Elections are meant to return a government and an opposition party/parties. Campaigns are meant to allow the public a chance to decide who should occupy those offices. As institutions, both are essential to democracy. The expectation is that voters collect information from the world—supplied by their friends and family, by professional organizations, by the media, by political parties and by other sources—develop attitudes and preferences, weigh their options and priorities, and then make a decision for whom to vote. However, the evidence as to whether or not voters can do this often or reliably is mixed. But rather than reviewing all the literature from studies of political behaviour—though I will look at some—my interest in this section is to examine how the *institutions* of campaigns and elections contribute to this mixed status.

What is it, then, about campaigns and elections, as institutions, that interact with cognition to produce sub-optimal outcomes? My sense is that there are at least three factors that contribute to the problem: speed, volume, and frequency. By speed, I mean the pace at which elections and campaigns occur. Typically, these are fast-paced contests that do not encourage reflection, and, indeed, tend to make it difficult, since by the time one has started chewing their electoral fodder it is already time for a new bite.

By volume, I am referring to the sheer amount of information that contemporary campaigns generate. When the volume of information produced is combined with speed, elections churn out massive amounts of information that is processed by the media and distributed as it happens with up-to-the-minute coverage. This churn of data includes both substantive material, such as policy, and procedural material, such as the so-called “horse-race” elements of polls and gaffes and the ups and downs of the campaign trail.

Finally, by frequency I am referring to how often elections take place: typically, not very often. While voters may have three levels of elections to pay attention to—municipal, provincial/state, and federal—and may have intermittent elections within those levels (e.g. bi-annual congressional and senate elections, by-elections, and so on), they tend to pay most attention to the four or five-year national elections. The relatively low frequency of electoral practice (which is relatively low as a practice compared to, say, driving a car or practicing a sport or doing your job) means that citizens are not given much opportunity to “get good” at being effective when it comes electoral politics. Moreover, given that our political system tends to conceive of civic duty as mere voting, or perhaps serving on a jury, it is not as if citizens are getting much democratic practice between elections—that is when it comes to affecting change. Jacobs et al. (2009) do find that in the United States, citizens are *regularly* exchanging political information, engaging in formal and informal political debate, and talking about the issues of the day. That, however, has not been enough to keep the U.S. from becoming a de facto oligarchy *when it comes to policy outcomes*,<sup>58</sup> as Page and Gilens (2014) found in their study of American policies from 1982-2002. Whatever the *cause of* inequitable policy outcomes (for instance, lobbying efforts), electoral politics has not been enough to deliver equitable policy outcomes. So, even when there are basic, regular political discussions occurring, their effect on policy may be limited (even if they affect campaigns and elections)—which suggests, among other things, a critique of aggregative democracy as it relates to transforming political preferences into formal outcomes.

---

<sup>58</sup> Page and Gilens did not use the term “oligarchy.” However, their findings—that the vast majority of policy favoured and was determined by a small elite—are consistent with typical understandings of the term oligarchy.

So, what does it mean to citizenship that our elections and campaigns are marked by speed, volume, and low frequency? On balance, it means that the emergent environment during elections is probably not particularly conducive to generating good judgments and decisions. Why not? Because high speed, high volume, and low-frequency campaigns challenge both ontological security and meaning structures—so much flux, so little stability. They are built—or have evolved—in such a way that our cognitive limitations are severely susceptible to exploitation and/or that elections prime ontological *insecurity* by introducing uncertainty around a central feature of (democratic) society—who gets to govern and how they get to do so (i.e. the sorts of policies and laws they adopt). It should come as no surprise to a student of politics to learn that political parties are aware of this and benefit from it. And while individuals have *some control* over what they pay attention to day-to-day, the overall process of contemporary politics is outside of their control.

Given the whirl of campaigns and our bounded cognitive capacities, when it comes to elections it is not surprising that voters do not seem to adopt the Bayesian, or economic, model of voter behaviour. This model imagines the voters dispassionately collecting data calculating expected utility based on past party behaviour and expected future behaviour, and updating their preference in response to changes (Downs 1957). Instead, it seems that voters rely on less purely rational determinants, such as family (Berelson: 1954), existing partisanship (Campbell et al. 1960, Green et al. 2002), demographic cues (Cutler 2001), or affective affinity (Lodge and Taber 2014).

Does it matter that voters rely on these heuristics? Another way to ask this question is: Would voters, absent these heuristics, choose differently? And also: Would they choose better?

After all, there are good reasons to use heuristics—they are faster, easier, and, perhaps, the individual simply trusts the source. With heuristics, however, individuals are effectively (at least partially) delegating their choice. The problem here is that with full information, ample time, and the motivation to scrutinize their options, individuals might choose otherwise. In essence, given robust cognitive autonomy and the rational consideration of their options, an individual might well choose differently than they do based on sociological, partisan identity, demographic, or affective cues. Given this fact alone, it follows that individuals ought to have a reasonable chance of making a decision based on at least an approximation of conditions that would enable them to reflect critically on their options and *to be precisely aware of what is motivating them to choose the option they are choosing.*

The affective affinity determinant noted above is particularly important. As we have seen, it seems likely that many voters use a kind of on-line model of information processing based on heuristic cues within which they evaluate a candidate through developing a *feeling* about them or some position they have taken before storing that *affective* evaluation and then forgetting the specifics that generated it (Lodge et al. 1995, Lodge and Taber 2014). It is important to note that within this model, voters remain responsive to campaign messages (Lodge et al. 1995). However, such responsiveness does not account for the quality of information that voters receive or their ability to process it in a reliable and accurate way; or, to put it in terms related to this dissertation: to act on that information in a rational, autonomous way. Instead, the fog of war generated by the speed of the campaign, the volume of information it produces, and the relative infrequency of its occurrence, forces individuals into cognitively difficult situations that generate sub-optimal judgments and decisions. The result

echoes the argument made by Kuklinski and Quirk (2000), and explored in chapters one and three that citizens *tend to use* heuristics poorly: off the shelf, uncritically, and with missing key information—a practice that generates structural errors in judgment and decisions. Kuklinski and Quirk attribute the phenomenon to shortcomings in our evolutionary capacities and an emergent gap, similar to the one I discuss throughout this dissertation, between what is expected of citizens and what they tend to be able to deliver.

So, on balance, campaigns and elections, *as institutions*, seem to exacerbate existing cognitive limitations—at least as they are currently structured. I say “as institutions” because the phenomenon is part of their very structures. Of course, having no campaigns would be worse from the standpoint of cognitive autonomy (not to mention democratic legitimacy). The point, rather, is that campaigns and elections could be done differently, they could be carried out in such a way that minimized the effect of our cognitive limitations on producing judgments and decisions. One way to reform campaigns would be to limit the influence of money and to set up rules that ensured that the heuristics voters consulted were more reliable. Another would be to set up aggressive regulations around advertisements that limited what candidates could do with sounds and images. Another still would be to ensure that candidates or their representatives engaged in sustained exchanges with their constituents on substantive matters. But elections are often not run this way. And so, given how they are set up, these institutions seem to contribute to the general problem of sub-optimal outcomes that I have been discussing throughout this dissertation.

### *The media and attitude formation*

The media is an institution. It is governed by more-or-less regular practices and relatively tight, widely accepted rules. While changes in mass communication technology and social media have started to blur the lines between traditional and new media, the two have not yet merged or otherwise become indistinct. I am interested in the press as an institution here insofar as they impact how the attitudes and opinions of individuals are generated. Still more specifically, I am interested in how priming and framing affect people. This section is related to the above look at campaigns and elections, but distinct from it insofar as the media continues its political coverage outside of election periods. It is this long-term commitment to collecting and disseminating political information that I am interested in in this section, as well as the question of how the media shapes information for consumption, and how that intersects cognitively and behaviourally with the individuals who consume that information.

Of course, the media cannot help but prime and frame. Given scarcity—time, money, attention spans, access to information—choices must be made about what to cover, how to cover it, and how often it should be covered. Now, before moving on, it is important to make a distinction between priming and framing. While the two are sometimes used interchangeably, they are quite different. *Priming* is about how individuals are deliberately or inadvertently prepared to evaluate an issue through exposure to external stimuli. As Iyengar and Kinder (1987) put it, priming “changes...the standards that people use to make political evaluations” (63); or, as Scheufele and Tewksbury (2007) put it, “Priming occurs when news content suggests to news audiences that they ought to use specific issues as benchmarks for evaluating the performance of leaders and governments” (11).

For instance, a story that appears more often or early in a newscast might be judged as more important (Iyengar et al. 1982). The cognitive mechanism behind priming is accessibility (Tversky and Kahneman 1973). So, priming is about changes to *the criteria people use to make judgments based on the availability of information*. This phenomenon is not the same thing as framing. *Framing*, again drawing on Scheufele and Tewksbury (2007), “...is based on the assumption that how an issue is characterized in news reports can have an influence on how it is understood by audiences” (11). Presenting information in different ways has different effects on how individuals. So, framing is about *how the information is structured* and rather than being cognitively driven by accessibility—i.e. via a memory-based system—it is about immediate evaluations based on forms of presentation.

For instance, as Kahneman (2014) has found in his research on prospect theory, individuals tend to be risk-averse and consequently sensitive to potential losses (see also Kahneman and Tversky 1974; Tversky and Kahneman 1981). The way that a problem is framed may thus have an impact on decisions around “risky” behaviour. In one study, medical students who were told that a new treatment had a “50% success rate” were more likely to think of it as a success and recommend it to their family than those who were told that the same treatment had a “50% failure rate” (Levin, Schnittjer, and Thee 1988).

Insofar as the media—the prime source for the dissemination of political information—must make choices about the information they present—what to present, how to present it, and how often—they are engaged in both priming and framing. Of course, this may be more deliberate and extensive (say, Fox News) or less (say, the BBC), but each will, nonetheless, occur. However, as I have argued in previous chapters, to the extent that priming and framing

move together structurally, and to the extent that they influence individuals disproportionately, they may contribute to people producing judgments and decisions that are less autonomous and rational than they might otherwise be. As Iyengar et al. (1982) found, counter-arguing reduces some of these effects, but only among “the politically expert.” For “the politically naïve,” media effects were much stronger.

What is troubling is that while framing and priming effects are unavoidable—in the media, in campaigns, even in deliberation—not enough attention seems to be paid to their effects by those disseminating messages, unless those messages are deliberately constructed to mislead. Because individuals are so susceptible to framing and priming, because autonomy is valued, and because both democratic legitimacy and substantive outcomes depend in part on the decisions of individuals who are receiving these messages, the stakes are high. Thus, we must ask: What can we do about it? Some research suggests that framing effects can be mitigated—or eliminated—when individuals have access to credible sources of advice (Druckman 2001). Moreover, “contextual forces” and expertise can also reduce framing effects—Druckman (2004) finds that elite competition, deliberation, and issue expertise help individuals to resist or overcome such effects. These findings imply that context and individual competence matter a great deal when it comes to ensuring that individuals are not being led astray by how a message is framed.

### *Political parties and partisanship*

Political parties are essential to contemporary representative democracy. The functions of parties are varied, and their benefits are many. As institutions, they aggregate interests, form

governments and oppositions, and, most importantly for my purposes in this dissertation, provide individuals with a structured and more-or-less reliable heuristic to guide their vote choice and to help them generate opinions and positions on issues. In this sense, parties act as meaning structures that allow individuals to develop a sense of ontological security around the electoral process. Parties, though they change through leadership and member shifts, and policy shifts, tend to generate enough stability to be reliable heuristics for voters over time. However, as institutions, parties can also lead to entrenched patterns of non-thought among supporters and can, it seems, even act as information-distorting lenses. Insofar as this is true, parties might then reduce autonomous, rational judgment and decision making by setting up conditions under which information is cognitively (structurally) distorted.

What happens when a heuristic enables—or outright becomes a tool for generating—bias? Bartels (2002) finds that long-term partisanship—understood as “party identification”—*biases perceptions* of information experienced by those who have longstanding party identification. He finds that it is not partisan *values*, but actual different experiences with the *information* itself. Partisan identification is, for some, a lens through which factual information is viewed, and such individuals adjust their *perception* (not their evaluation) of the information, non-consciously, in an attempt to maintain a favourable assessment of their party of choice. This, in the context of the United States, concludes Bartels, helps to keep distance between the Democrats and Republicans and their supporters. Evans and Pickup (2010) support Bartels’ argument by looking at retrospective voting in United States presidential elections from 2000-2004. They find that political preferences—linked to parties—shape economic perceptions, rather than vice versa. Or, as they put it, partisanship mediates economic

perceptions/presidential support affects economic evaluations. In this case, the perception of *economic facts* is viewed and shaped through a partisan lens.

At this point, parties become more than a heuristic. This seems to occur through an individual's *identification* with a party—an identification that forms early in life, persists, and determines how individuals *perceive* of political phenomena (Green et al. 2002). Of course, occurrences of information distortion may not occur for all of those who use parties as heuristics (or, if they do, may not occur regularly), but they are most likely most prevalent among those who expressly identify with parties. Indeed, Lenz (2012) finds that partisans tend to “blindly” support politicians they like or prefer (although this is not because of *policy positions* held by either); and while support is partially contingent on performance, including perceptions of the politician's competence or honesty, other considerations include physical appearance and, as noted, party identification. According to Lenz's findings, policy has little to do with a voter's support for a politician, even filtered through the lens of partisanship—individuals merely “follow the lead” of their preferred politician. In the same vein, Levendusky (2009) finds that partisan identity is stable and individuals use that identity to sort their issue stances (rather than using policy to sort their party or vote preferences); however, he also finds that voters will on occasion switch their party identification based on the issues (although this is less common than the converse).

In the case of partisanship, attempts at making sense of the world seem to involve at least some non-conscious distortion of information, aimed at maintaining certain meaning frameworks. If my party is the party of fiscal responsibility, and if they form government and subsequently spend heavily during a certain period, then I have some work to do; I can either

change my evaluation of them, thus upsetting my understanding of my party identification and of the political world, or I can adjust my perception of the information itself—again, outside of my awareness. It seems that for some who have strong party identification, it is the latter approach that is adopted.

This phenomenon reduces politics in some instances to following the pied piper. In contrast with both a Bayesian updating model (Downs 1957) and a practical wisdom model (Cameron unfinished manuscript, Sharpe and Schwarz 2010), there is little calculation or problem solving occurring outside of drawing on long-established patterns of expectation and belief to guide current evaluations through the non-conscious and distorted perception of facts and arguments. What is problematic about this phenomenon is that individuals, had their perception not been severely biased, might come to different conclusions than they ones they reach. So, parties, as institutions with long-standing practices and ideological tendencies, seem to generate stability at the cost of constant readjustment of reality in the direction of one's pre-existing ideological or partisan preference.

So what? What's to blame for these shortcomings? An initial assessment indicates that both institutions and cognitive tendencies are to blame for autonomy and rationality being undermined in the case of partisanship. (After all, if there were no parties, partisan identity would likely operate very differently—or not at all.) It is outside the scope of this dissertation to get deep into the question of how we might *directly* improve these institutions themselves, but it is worth noting that changes to this—and other institutions—might enhance judgment and decision-making *outside of* democratic deliberation. While my focus is on enhancing judgment and decision-making internally deliberation, if parties could be made to act as mobilization and

policy/preference-aggregation institutions rather than as cognitive lenses, we might expect to see improvements in the quality of thinking of certain partisans.

### *Electoral systems*

The primary purpose of an electoral system is to allow citizens to elect a government to act on their behalf. Different systems will tend to return different types of governments and, depending on the political history and political culture of a given country, state/province, or municipality, may tend to favour certain parties or kinds of parties over others. First-past-the-post (FPTP) systems may encourage strategic thinking by voters (who may vote for a non-first choice party to keep out a party they prefer even less), tend to produce a two-party system, and tend to marginalize parties with less or more-dispersed popular support; proportional systems open up more space for parties that do not enjoy broad or regionally concentrated support (Duverger 1951, Johnston 2008; cf. Riker 1982).<sup>59</sup>

In this section, I will focus specifically, though briefly, on the question of how electoral rules around voting impact individuals. In a 2012 experiment surrounding the French presidential election, Van der Straaten et al. (2013) found that a significant minority (10 to 15 percent) voted differently under alternative systems. This finding—which, the Van de Straaten et al. note is consistent with the literature on vote choice in alternative electoral systems—suggests that the act of voting is, for some at least, strategic and malleable. Thus, electoral institutions shape the considerations of voters in such a way that those brought to bear on vote

---

<sup>59</sup> There are, however, important exceptions or variations, such as India, which uses FPTP but has several parties represented in Parliament and Canada, which has averaged nearly five parties since the 1930s

choice are at least somewhat complex—independent of the question of whether or not voters are receiving good information and processing it accurately (with reference to factual information), a concern addressed above. So, the *type of electoral system* clearly matters when it comes to considerations of who someone is going to vote for and, consequently, both the individual candidates who are returned and the government that is formed in an election.

For instance, in a FPTP system, where an individual or party needs only a plurality of votes to win, the strategic considerations by both parties and individual voters—such as how to frame the party and election, the sorts of policies that will appear in a platform, and so on—will be different than in, say, a majoritarian (e.g. a ranked-ballot) system in which the support of 50 per cent plus one is required. The implications are that electoral systems will generate different incentives for parties and candidates in regards to both *how* to frame the information they provide and *to whom* that information will be directed. For example, if Party A is conservative, and they expect that 40 per cent of likely voters are conservative or near-conservative, in a FPTP *where there are more than two competitive parties*, there will be a strong incentive for the conservative party to frame their message in ways that will appeal to approximately 40 per cent of voters—those who conservative/near-conservative. In a majoritarian system, one would expect there to be more convergence towards the political centre where parties would battle over both liberal/progressive voters and more centrist/centre-right voters. But what does this have to do with how individuals think about politics? For one, how a party frames its messages will affect how voters think about and process those messages. This practice serves to remind us that there is no neutral way to engage in political information processing—it thus remains up

to the individual to be aware of where parties are coming from in the context of an election (and an electoral system) as they relate to the mechanical element of the electoral return.

Where things get even more interesting is when Duverger's (1951) concept of the "mechanical" effect is contrasted with his concept of the "psychological" effect in the context of producing an electoral return. Here, two elements of the institutional set up of an electoral system meet and, indeed, run up against one another. While the former refers to how votes are counted—*after* voting has taken place—the latter refers to how individuals choose who to vote for *before* they vote (Blais et al. 2012). The findings from a study (i.e. an election experiment undertaken during an actual election campaign, in this case the 2011 Ontario provincial election) by Blais et al. (2012) confirm that the magnitude of psychological versus mechanical effects during an election, and whether or not they diverge or converge, is at least partly contingent on whether the electoral system is FPTP, proportional, or alternative vote. They also find that the relative impact of the effects varies *depending on which party you consider*. In another experiment (Van de Straaten et al. 2010) found that under certain (constrained) electoral conditions, rational choice models were predictive of behaviour: that is, under limited-complexity scenarios, at least when it came to strategic voting behaviour, through heuristics and "sincere" voting. Again, though, this depends on the level of required complexity being low. The main takeaway from these findings—unsurprising as they are—is that institutional set ups matter, at least to some degree, when it comes to *how* individuals think about voting. The findings discussed above also reveal a tension between changes in electoral outcomes based on different institutional set-ups and the quality of decisions made as information and rule complexity increase.

What does this mean for judgment and decision making? It depends on what sorts of outcomes one wants—and whether they prefer a focus on policy, strategy, or some other consideration when it comes to elections. From the perspective of someone who wishes to see more autonomy and rationality built-into electoral decisions, the key is less about the specific system that is chosen and more about preparing citizens to interact with the particularities of whichever system that happens to be. Since different institutional arrangements will produce different incentives, strategies, and tactics for parties, they will require different approaches for voters. However, regardless of which system is chosen, one take-away is that increasing complexity seems to undermine strictly-rational choices—a reminder that at least in the context of an election, when it comes to complexity, after a certain amount, less is more.

*Ideology and hegemony: a special case*

I have put ideology (as an organizing set of beliefs) and hegemony (in the Gramscian sense of a "cultural leadership") together in this section, though I consider only the latter to be an institution in the sense I refer to it here. However, it is important to examine the former alongside hegemony, which I will discuss below. Ideology relates to hegemony and informs it in some ways, but particular ideologies flow from hegemony, or, at least, are enabled and constrained by it. Accordingly, hegemony is the institution here, while ideology, in the sense I mean it, refers to a series of political beliefs set that is shaped within—and which also shapes in return—hegemony. Ideology and hegemony are a special case because they are, together, powerful determinants of the aforementioned political institutions, enabling and shaping them through providing a series of norms, values, concepts, and ideas that are informally sanctioned

as legitimate and placed on the political and cultural register of acceptable ways of thinking about, talking about, and practicing politics. I will discuss each in turn before relating them to one another in particular, and to institutions in general.

First, ideology. For this section, I take ideology to be a “set of beliefs about the proper order of society and how it can be achieved” (Erikson and Tedin 2003: 64). I have chosen to disaggregate ideology from parties and partisanship, but, obviously, it connects to both. In a way, I think of partisanship as applied ideology, though I can imagine contexts in which this is not true (e.g. in cases of those who are after patronage or pursuing blatant self-interest)—though one would expect that given the incompatibility of ideology and party, something would have to give in one direction or another before too long. Of course, ideology could be framed in other ways than as a tool for institutional hegemony; indeed, it can be seen as a perspective or a commitment, or even, like partisanship, a lens that shapes our view of the world (not that institutions are necessarily ideational). In a sense, this latter point is what interests me about ideology and why I include it as an institution.

More specifically, I am casting ideology for the purposes of this dissertation as expressly political: ideology, then, is also a set of patterned rules for *deciding what counts as what in a particular political context*. Is the act of terminating a pregnancy the expression of a right or the committing of a moral (or legal) wrong? Is shouting a hateful epithet an instance of free speech or criminal abuse? Is waging war the duty of a state to keep its people secure or is it the abandonment of that duty? Ideology, for *some people*, structures and patterns responses to these questions. And insofar as ideology does this, and in so doing creates a theoretically strong and stable point of reference for judgments and decisions, I see it as being, among other things,

a key applied outgrowth of hegemony. So, how does ideology intersect with cognition, judgment, and decision making?

Converse (1964) finds that the nature of mass belief systems (when it comes to specific policy issues) is such that very few individuals remain consistent or stable in their views over time (although their *professed* ideology might remain stable). Rather, he finds a high level of randomness and mere group interest among non-elites (see also Jennings 1992; cf. Freeze and Montgomery 2016). Nonetheless, ideology seems to be a factor in the political lives of many (and sometimes a factor *outside* those lives, too). Data from the Canadian experience shows that those who find themselves on the left and right of the political spectrum organize their opinions differently, with those on the left demonstrating more ideological coherence than those on the right (Belanger and Stephenson 2010). This results in those on the left tending to have more coherent and ideological consistency than those on the right. Taking ideology as an organizing institution, then, we should expect more predictable judgments from those on the left, though this might come at the expense of being able to adjust to a variety of complicated social and political contexts; however, research on this is limited, and we may have good reason to doubt just how appropriate the left-right distinction is for research (and perhaps more broadly), since individuals think of left-right in (sometimes significantly) different ways (Bauer et al. 2016). Paxton and Kunovich (2000) find that ideology even affects the number of women who are elected to national legislatures through different perspectives on the role and place of women in social and political lives. And, of course, ideology has been shown to act as a stabilizing and sense-making tool (e.g. by providing explanations for political and social

phenomena, and by providing existential security about one's place in the world)—especially on the right (Ball and Dagger 1992: 1-2, Jost et al. 2003, Jost et al. 2008).

But if ideology acts as a sense-making tool (within hegemony), specifically related to political dispositions, it does so, as I have suggested, at the cost of smoothing over the rugged complexities of social and political life and the competing values, preferences, and needs of those with whom one disagrees. Ideology (the ideational component of partisanship for the very few who consistently think ideologically) is thus a lens that shapes one's perception of the world as it enables one to better make sense of the world. This, of course, makes it a useful tool for an aggregative approach to democracy—since such a system is largely about competition among existing preferences—but a rather poor and even counterproductive one for a deliberative system aimed at generating the epistemically best judgments possible through a collective process of reason giving. One of the reasons—perhaps the most significant—why ideology is counter-productive when it comes to reaching the deliberative goals mentioned is that it seems to discourage the kind of careful and critical reasoning discussed earlier and required to kick an individual out of 'reasoning auto-pilot.' Insofar as ideology (like partisanship) offers a pre-packaged program rather than a critical filter, it risks undermining one of the key requirements for good democratic deliberation: openness.

Next, hegemony. When I refer to hegemony here, I am thinking of it the sense of cultural hegemony as it is conceived by Gramsci (1992). Hegemony in the Gramscian sense refers to “the cultural leadership of the dominant classes in the production of generalized meanings, of 'spontaneous' consent to the prevailing arrangement of social relations—a process, however, that is never finished because hegemony can never be complete” (Ang

1996). Hegemony is a force and a milieu—it refers to a series of generally consistent and compatible norms, values, concepts, and ideas that shape behaviour and inform what most people tend to think of as common sense and consensus. Hegemony generates mass consent (manufactures it, as Gramsci would put it) for dominant political and cultural practices.

For better or worse—my intention here is not to engage in a debate on Marxian philosophy or dialectics—hegemony is a macro- or meta-institution that shapes ways of thinking about the world, and, consequently, ways of behaving in the world. (For Gramsci, ideology-embedded institutions make up hegemony). Through hegemony, the incentives and disincentives of certain institutions, as we have seen in this chapter, are enabled. For example, in the United States, mass media coverage of political scandals and “scandals” is pervasive and provides Americans with a common topic of political conversation (the political equivalent of “How about this weather?”). Regardless of the nature and seriousness of the scandal/“scandal,” a combination of incentives generated by overlapping institutions (e.g. the marketplace through ratings and advertisements, the media itself through the need to justify the profession and to “break” news, etc.) encourage media outlets to cover the topic and individuals to discuss it (e.g. to generate a sense of belonging and in-group solidarity). However, while coverage of and discussions about scandals occur regularly, there has been no challenge to the nature of the political system (and its institutions and practices) that enable and encourage malfeasance in the first place. The system and its institutions are protected under the banner of hegemony while “a few bad apples” are dismissed as outliers in need of punishment and perhaps rehabilitation. The same analysis could easily and extensively be applied to corporate America many times over.

Again, ideologies are sets of narratives for organizing and orienting people in the world; they intersect with hegemony as an institution insofar as they emerge from it and feed back into it. For instance, if liberalism as an ideology privileges individual (negative) freedom and self-ownership as natural and morally right, then capitalism as an economic system rests comfortably (and perhaps solely) on the register of acceptable and appropriate economic systems. Particular ideologies may then be informed by this reality—though to different degrees. Take, for instance, the Canadian political spectrum, where even the leftist/“socialist” New Democratic Party (who, incidentally, dropped the word “socialist” from the preamble to their party constitution in 2013)<sup>60</sup> are willing to embrace free markets and typical capitalist logic. The NDP is an organization that originally began as a radical, prairie socialist alternative (at the time known as the Co-Operative Commonwealth Federation and the Canadian Labour Congress) to the big two establishment parties in Canada: The Liberal Party and the Conservative Party. They have since moved towards the ideological centre—a move that reflects adherence to hegemony and which further legitimizes hegemony as an organizing institution. The ideological bandwidth in Canada is narrow—which is not to say that it does not matter to or for Canadians which party governs—reflecting a broad consensus consistent with cultural hegemony as a key organizing institution.

So, why do hegemony and ideology matter for cognition? Together they provide a powerful structure for thinking about the world—whether one thinks in a consistently ideological way or not. Thus, to understand political thinking, we must understand the

---

<sup>60</sup> See CBC, 2013: <http://www.cbc.ca/news/politics/ndp-votes-to-take-socialism-out-of-party-constitution-1.1385171>.

narratives that guide that thinking; when it comes to democratic deliberation and attempts to enhance autonomy and rationality, it is important that individuals are given the opportunity—and tools necessary—for challenging and thinking outside of this structure. To the extent that ideology and hegemony are *merely replicated* without critical reflection—perhaps as a variation on a theme—in a deliberation, or elsewhere, individual autonomy and rationality are undermined, as (perhaps) is the key deliberative requirement: reason-giving. To the extent that individual thinking is dominated by pre-arranged or rigidly patterned sets of scripts, that thinking cannot be deliberative.

## **Conclusion and summary**

The political institutions that collectively help determine how we live our political lives are not neutral from a judgment and decision-making point of view. While they enable and constrain behaviour and provide for the necessary psychological foundations for political—and other forms of—life, they do not merely set neutral parameters within which individuals create the content of their political world. To a significant degree, *institutions are our political world*.

Different institutional set-ups will create different incentives, sanctions, and will encourage or shape different ways of thinking and judging political matters. Some institutions will encourage better thinking than others; some will encourage more representative outcomes, too, whether regarding electoral returns, policies, or, more broadly, inclusiveness.

It is important that we further study how institutions interact with human cognition and what that means for individual judgment and collective decision making, whether this is understood in terms of practical wisdom rationality or something more akin to rational choice

understandings of human behaviour—each of which has a particular place when it comes to understanding human behaviour both in general and within the context of political, institutional life. The more we learn about this, the better equipped we will be to choose and design institutions that maximize the quality of political outputs while generating better political agents through both increase capacity and institutional responsiveness. And just as importantly, in understanding how institutions shape thinking, we can better adopt a critical lens to interrogate how they enable, constrain, and otherwise shape not only *how* we think, but also *the sorts of things we are likely or able to conceive of/think about in the first place* (e.g. in the instance of hegemony).

## Chapter 7: A deliberative systems approach to generating epistemically better democratic deliberation

Throughout this dissertation I have argued that there is a gulf between what proponents of deliberative democracy normatively *expect* from citizens who engage in public, political deliberation and what many of those citizens are capable of *delivering*. I have argued that this gap is, in part, due to the fact that while these proponents expect autonomous, rational deliberators who bring to the table publicly justifiable reasons for and against certain preferences, our evolved cognitive capacities are not always so well-suited to generating such reasons autonomously and rationally, which undermines the epistemic defense of democratic deliberation by undercutting attempts at establishing epistemic validity through reasoning-giving, judgment, and decision making. Moreover, certain of our social and political institutions *work against us* in our striving for more autonomous and rational judgments and decisions, before, during, and after deliberation while they should be helping to correct weaknesses and amplify our capacities. Accordingly, if we are to generate better outcomes—both in terms of personal judgments and collective decisions—we need to work on both our personal capacities and practices *and* the social and political institutions that surround us to bring them closer into line with what we expect from one another as engaged citizens capable of self-government.

This chapter is concerned with examining the concept of a deliberative system—that is, deliberative democracy scaled up from individual, occasional, or one-off deliberative events and entrenched as part of a broader, long-term structure of institutions. Such a system includes several components with which deliberative practices will interact, such as parliament, protests, news media, and advertising. In this chapter I do three things: first, I will explain what a deliberative system is beginning with the landmark 2012 volume *Deliberative Systems* and

then presenting other approaches. I will outline the elements and functions of such a (largely still-theoretical) system. Second, I will pay specific attention to the epistemic function of a deliberative system and the role of experts within the system. And finally, I will argue that a deliberative system, properly constituted, arranged, and scaled, can form an important, central part of the solution to the problem closing the gap between what we expect from citizen deliberators and what they tend to be able to deliver—without sacrificing our more ambitious normative expectations. Chapter five of this dissertation is about how certain specific personal practices and individual institutional designs could generate better judgments and decisions—this chapter complements it with a discussion of how those practices and individual institutional designs can be corralled and put to work in the name of deliberative democracy in contemporary mass democracies.

### **Understanding the deliberative system: Concerns, functions, and roles**

In this section, I will explain what precisely the concept of a deliberative system entails in theory and practice, and outline what such a system is *concerned with*, which broader *functions it serves*, and which *roles experts and citizens play* in a democracy. The study of deliberative systems is a fairly recent development—the most serious research began in 2012—and so there remains much to be worked out both theoretically and empirically, especially when it comes to understanding how a deliberative system might scale. Nonetheless, a basic conceptual scaffolding of the deliberative system is available.

*What are the concerns of a deliberative system?*

Jane Mansbridge and her colleagues (Mansbridge et al. 2012) think of deliberative system analysis as regarding three core *concerns*: first, that deliberation is considered “as a whole” and in “large scale societal terms.” In this way, deliberative democracy is considered on a larger scale (2-3). They also conceive of such analysis as looking into the “division of labour among parts of a system” (3). Finally, they argue that deliberative systems analysis allows for an examination of “large contextual issues and broad systematic inadequacies that have an impact on individual sites and shape the possibilities of effective deliberation” (3-4). So, when speaking of a deliberative system, as opposed to a deliberative event, the critical distinction is that the systems approach is nearly-all-encompassing of the elements of mass democratic governance, both inside and outside of deliberation itself. And, of course, the span of time is expected, in the case of a deliberative system, to be long-term, whereas any individual deliberative practice could be semi-permanent, rare, or even a one-off event devised for some specific, extraordinary need.

### *What are the functions of a deliberative system?*

Within their analysis of the particular effects of a deliberative system, Mansbridge et al. include three *core functions*, which roughly correspond to the functions of deliberation outlined in chapter one: an *epistemic* function, which is concerned with producing “preferences, opinions, and decisions that are appropriately informed by facts and logic and are the outcome of substantive and meaningful consideration of relevant reasons” (11); an *ethical* function, which is about promoting “mutual respect among citizens” (11) through listening and reason giving and respecting divergent preferences; and, finally, a *democratic* function, which is about

including “multiple and plural voices, interests, concerns, and claims on the basis of feasible equality” (12). These functions of a deliberative system are meant to exist, entrenched, within the broader system and are meant to guide the overall functioning of a democratic order. However, they are not meant to be *entirely pervasive requirements*—which is to say that within a deliberative *system*, according to Mansbridge et al., there is space for certain elements that are in tension with these functions, sometimes complementing them, sometimes challenging them. These exceptions include, among other things, space for experts, pressure groups, protests, and partisan (and other) political media. One of the strengths of a deliberative system is that elements have complementary roles to play, including some that are not, strictly speaking, themselves democratic, but which are checked and balanced against others that are more democratic so that, on balance, the *system itself* reflects the values, priorities, and practices required for a system to be democratic.

Now, in this dissertation, I am primarily concerned with the *epistemic function* of deliberation. Therefore, two parts of the literature on deliberative systems stand out: the specific epistemic function of the system—which is a kind of scaled-up version of the epistemic function of individual deliberations, outlined by Warren (2002), which I examined in chapters one through four—and more specifically the role of *experts* within the system. (However, as I will note in this chapter, the other sorts of functions of a deliberative system—ethical and democratic—are in some ways related to epistemic goods, too). In the case of the epistemic function of a deliberative system, my interest comes from the overarching question of this dissertation: Can deliberation generate good epistemic judgments and decisions? In the case of a deliberative *system*, this question cannot be answered by looking at any individual kind of

deliberation or deliberative moment. Instead, it must be explained concerning the *whole of the system*. While answering that question is work for another project—since this one is concerned with deliberations themselves—it is still worth looking at the relevant individual elements of the system since a systems approach allows for a different and broader perspective for examining deliberation as such. In the case of a deliberative system and its epistemic function at large, it makes it possible to ask what the role of experts and citizens might be *writ large* and how those two functions might—or might not—lead to improving the epistemic quality of deliberation as such. (Hence my interest in deliberative systems, their epistemic functions, and the role of experts within them for this dissertation.)

But before moving on it is important to make clear a point about *where deliberation itself should occur within the system*. We have briefly looked at the functions and elements of deliberation, and a bit at how they intersect, but within a deliberative system itself where would the deliberations take place? Currently, sites of deliberation vary and within a deliberative system particular setups could also vary: that is, each might have a different distribution of deliberative and non-deliberative bodies as well as different deliberative and other decision-making procedures. For instance, Parkinson's examination of deliberative systems (2006) focuses most closely deliberative juries and deliberative polls; Dryzek (2009) notes that "legislatures, cabinets, constitutional courts, and corporatist councils that empower representatives of labor and business federations and government executives" (1382-1383)—the "central institutions of states"—can host deliberation. He also notes that Rawls considered the United States Supreme Court to be a deliberative body and he also includes various kinds of forums—the sort regularly thought of as deliberative, such as assemblies and stakeholder

dialogues—as deliberative spaces (1383). While Dryzek echoes the point raised by Mansbridge et al. above, that different elements of the system contribute to outcomes in general and deliberative outcomes in particular, and some can make up for dysfunction in others (Dryzek 2009: 1383, 1388).

Jürgen Habermas is one of the strongest proponents of public sphere deliberation, which includes, among other sorts of engagement, face-to-face deliberation. He separates deliberation into a larger public sphere of less structured political talk and a smaller formal, deliberative sphere of representatives so that deficiencies in public deliberation possibilities can be addressed (Habermas 1996: 182; see also Benhabib 1996 and Goodin 2005: 182). Hendriks (2006) takes a similar approach, imagining inclusive deliberation “occurring in a range of discursive spheres” (486)—transcending the “micro” and “macro” deliberation model that locate deliberation, in the case of the former, among civil society actors in structured moments and, in the latter, an unstructured public at a large (487). But, wherever deliberation occurs, what is most important in a *democratic* deliberative system is having sufficient open, inclusive, and responsive spaces for deliberation that connect citizens to their representatives and, ultimately, to the collective decisions that affect their lives. Nonetheless, with that in mind, I hasten to add that my central interest in this dissertation is *democratic deliberation undertaken by ordinary citizens* (who are often supported by and interact with experts and policy-makers); when I examine factors outside that which affect deliberation, including where and how deliberations occur, it is primarily due to their effects on the epistemic functions and force of deliberation itself rather than any ethical or democratic effects that might emerge—as we will see in the pages that follow.

*What are the roles of experts, policy-makers, and citizens in a deliberative system?*

Christiano (2012) argues that the purpose of deliberative democracy is twofold: first, epistemically, to generate shared citizen understanding in a political context and second, practically, to justly represent and disseminate the interests of citizens with an aim to creating policies and laws appropriate for that polity within the boundaries of justice (27). However, given the necessity and the benefits of specialization in contemporary mass societies, not all individuals of a given polity have the same role within a deliberative system: there must be a division of labour between politicians, experts, and ordinary citizens (27-31). After all, no one can do everything, and so-called “ordinary citizens” have lives outside of politics, which means that they have, at best, limited time for political matters and less opportunity to specialize in the many policy areas that a contemporary mass democracy must consider. Politicians, bureaucrats, and experts, however, by nature of their professions should have the necessary expertise and time to engage in the examination, creation, and implementation of complex laws and policies. The key, according to Christiano, is getting the division of labour right such that the resulting setup is just and democratic and also functional and effective—which requires that each part of the division works to enable the others to function properly (31-32).

According to Christiano, the role of citizens in a deliberative system, generally, is to decide the “basic aims” that a society is to pursue; they do this by electing representatives with different “packages of aims” and lawmakers who then, within the legislature, negotiate among one another—alongside non-politician experts, who have a role in helping generate information required to make policy decisions, to decide which packages, or elements thereof, will be

pursued, when, and how (33). The executive politicians and administrators then determine how laws and policies will be implemented (34). Citizens then play an “indirect role” in evaluating the implementation of these policies and laws. Citizens are thus generalists who act as broad agenda-setters as well as checks against executive and legislative politicians (who are sometimes, depending on the political system in question, one and the same). Alongside being responsible for setting the general direction of state affairs, citizens, according to Christiano, should also be “sources of different and competing research programmes in the various expert domains” (including the aims of expert research) and “evaluators of the pursuit of aims to whom the rest of society is accountable” (Christiano 2012: 36, 47). In essence, citizens act as general agenda-setters, sources of approaches to fulfilling the pursuit of that agenda, and checks against those who are tasked with the job of determining and implementing the particular laws and policies that are most likely to achieve the broad aims set out by citizens while also negotiating their many, varied, and often conflicting preferences.

As you can see, in this model, citizens set the basic aims of the democratic agenda, experts inform citizens and policy-makers on areas in which technical direction is required, and policy-makers make (most) final decisions (and are accountable to citizens through elections and, in some case, recall mechanisms). Together, these relationships are meant to form a symbiotic relationship that produces democratically acceptable and correct outcomes. According to Christiano, in this deliberative system design, “the deliberations are started by one group and completed by another” (Christiano 2012: 51). And within it, each deliberative sphere is, to some degree, checked by the other: the agenda set by citizens is checked by experts who evaluate outputs (though citizens may certainly override experts) and policy-makers (who may

override either/or both experts and citizens) who must negotiate any disagreements and divergences between experts and citizens, and among or between citizens themselves; the work of experts is checked by citizens who may choose, en masse, to reject their conclusions; and, finally, deliberations and subsequent decisions among politicians may be affirmed or corrected by citizens at the ballot box (or through other democratic mechanisms such as recall or ballot initiative, should they be available). So, these interwoven relationships within a deliberative system (including politicians as experts) demonstrates generally one element of a deliberative system that is meant to generate better decisions through three mechanisms: first, an efficient and effective division of labour (as we have seen above); second, a system of checks and balances within which each of the groups within the division of labour are subject to some constraint; and, third, increased concern with and attendance to truth sensitivity. The first two points have been discussed above, but it is the third point, truth sensitivity, that concerns me most insofar as it relates most closely to the matters I have been concerned with in this dissertation.

According to Christiano, “truth sensitivity” refers to a decision process that “does not ignore the best available science” and that “includes a method for assessing the community of knowledge itself” so that the process is not hijacked by special interests (Christiano 2012: 48-49). We have already seen, just above, a glimpse at what this requires: namely, an entrenched and vibrant democratic pluralism that generates substantive debate and multiple theories that must compete in public. As Christiano argues, “Without this kind of robust debate, parochialism, group-think, and cognitive bias will distort the process of the production of knowledge that will defeat its truth-sensitivity” (49). So, democracy becomes a “necessary

condition for the truth sensitivity of the community of experts...” in a “competitive struggle of ideas that can ensure that the process of social science is responsive to a lot of different sources of evidence” (49). (However, I would argue that at the level of the electoral system, democracy is not necessarily a *sufficient* condition for truth-sensitivity; see, for instance, the 2016 United States general election). I would broaden things by suggesting that more than “social science” should be included in this category: indeed, any science or knowledge-producing and sorting endeavour that is of sufficient public concern could be included here. But Christiano’s core argument stands: democratic pluralism serves as a tool that enhances truth sensitivity *in deliberation* by multiplying and amplifying voices within the debate and helping to prevent, both before and after the fact, the few and most vocal perspectives from dominating deliberations within a *deliberative system*.

That final point—that this pluralism plays out within a deliberative system—is an essential one. After all, any one-off deliberation might be hijacked by special interests or might, through randomness, fail to include a sufficiently broad series of perspectives, and, occasionally, some particular element of a deliberative system might become subject to an unrepresentative clique. But, *on aggregate and within the deliberative system*, in the long run, such advantages should be (that is, normatively speaking) washed out, limiting any structural bias within the system that unfairly privileges some group or groups over others. The trick, of course, is generating approaches that will achieve this outcome. In this dissertation, I have offered some approaches that, I argue, ought to be tested towards this end.

Imagine, for instance, a deliberation held right after a tragedy—whether it is an environmental catastrophe, a terrorist attack, or some other such event that increases the

salience of some perspective that may or may not reflect broader patterns, concerns, or risks. While such an event might be evidence in a broader discussion, its effect on the immediate perspective and approach of those who deliberate will matter—and not necessarily for the better. Calls for limit or to outright ban refugee seekers from entering a country after a terrorist attack comes to mind as a recent, regrettable, example of this. Even in a deliberation, it is hard to imagine shielding participants from the very strong force of the affective appeal to such kinds of arguments following events that disrupt ontological security, induce fear, and encourage fast, unreflective, visceral thinking—hence the need for time and repetition when holding deliberations.

### **Deliberative systems and cognitive error**

Is the existence of robust and varied elements, groups, and practices within a deliberative system alone enough to address the challenges to autonomy and rationality that I have discussed throughout this dissertation? In practice, to date, we cannot be sure, though theoretically a deliberative systems approach holds promise. Christiano notes that the ideals of deliberative systems—and I would add, to no surprise I am sure, the ideals of deliberation itself—are far from being achieved (49-50). Nonetheless, we have good enough reason to generate a robust and extensive theory of a deliberative system or systems worthy of testing in both experiments and day-to-day life. This “good enough reason” comes from a surfeit of data related the effects of institutions on behaviour (see chapter six), the effects of deliberation on judgment, and the effects of deliberative and institutional design on individual cognition. For instance, as noted earlier, research by Landemore (2013) and Page (2007) suggest that

cognitive diversity as epistemic benefits for deliberation and problem solving; building diversity into a democratic *system* allows for multiple sites of epistemic diversity that can work to generate better judgments and decisions—and gives us a functional reason, on top of pre-existing *ethical* reasons (i.e. the *ethical* function of a deliberative system), for supporting inclusiveness in democratic governance. For similar reasons, the *democratic* function of a deliberative system—mutual respect, reason-giving, and listening—also serves the *epistemic* function by enabling conditions for better judgments and decisions to be made, such as requiring that those who deliberate substantively engage in the process of justifying their preferences and accepting challenges from those who disagree with them.

It seems plausible, and indeed likely, that the epistemic (and ethical and democratic) goods claimed to emerge by such a type of system should be borne out, at least in part. A deliberative system offers the potential of serving as a site for the structural entrenching of the practices I discussed (conceptually) in chapter five. A systems approach allows for both individual, cognitive correctives to be practiced (through more-or-less standardized deliberative protocols across multiple sites of deliberation) *and* institutional correctives (through the establishment of the regular deliberative practices themselves, which will generate particular incentives, disincentives, and rules which, given proper institutional design and execution, we should expect to be of salutary effect on behaviour). So, in fact, a deliberative systems approach not only further and systematically addresses the particular cognitive biases and distortions I discussed in chapters three to five; it also tackles the challenges to good decision making and judgment generated by certain institutions, which was the focus of chapter six. Given this, in the following section, I will examine some of the effects of a deliberative system

and discuss their potential—or known—effect on deliberation, judgment, and decision making at the individual and institutional levels.

### **More deliberations, integrated deliberations**

A deliberative system can increase the number, frequency, and location of deliberative moments—e.g. citizen juries or deliberative polls—and thus provide opportunities for two key democratic benefits: more opportunities for practicing deliberation and a chance to represent a greater proportion of the population *across and within* deliberations, which, as I have noted, are important for deliberative skill building, preference stabilization and clarification, and diverse inclusion that produces more creative judgments and decisions. In regards to the *practice* as it relates to being a democratic good, my argument in favour of more deliberation(s) relies on the Aristotelean notion of practical wisdom insofar as deliberation becomes, in part, about the *practice of* democratic engagement: knowing *how to think about politics, how to understand others, and how to communicate your ideas* in a way that will be constructively received by others (though rhetoric is, of course, subject to abuse). In order to build these capacities, regular practice over time is required, which itself requires a number of deliberative opportunities. In regards to the latter, more deliberations mean that a greater proportion of the population will have an opportunity to participate and to be heard across and within deliberations (cf. Parkinson 2006: 149-150). By *across* I mean that a greater number of citizens will have the opportunity to engage in deliberations throughout their lives, while by *within* I mean that each instance of deliberation will be more likely to be marked by a diversity of citizens with different histories, ethnicities, preferences, political and religious affiliations, and

so forth. Ethical justifications aside, a commitment to more, more inclusive, and better-designed and structured deliberations are, I believe, likely to produce salutary effects that will enhance not just deliberative democracy, but democracy writ large as they are scaled up over time.

These benefits should produce further cognitively (good) benefits, as I argued in chapter five: that is, cognitive trends that help combat against a-rational pressures generated from small group-think, inertia, and a lack of citizen capacity due to low levels of practice. Of course, as I have noted throughout this dissertation, improving the epistemic reliability—i.e. generating good judgments and decisions—in a deliberative context requires that *several institutional and personal factors* are considered, and tactics, strategies, and designs are adopted. It may well be the case—indeed I am fairly confident that it *is* the case—that simply having more opportunities for more citizens to deliberate will not solve the challenges to achieving good deliberation I have discussed in this dissertation. Nonetheless, the increased frequency of deliberative opportunities—through a deliberative system—is a significant *part of* the solution. Indeed, as Brady et al. (1995) find in their resource model of political participation, civic skills are an important requirement to being able to engage in civic life—skills that are acquired early in life, often *outside of politics*. The researchers argue that these skills are developed at work, through participation in organizations, and even through engagement in shared religious practice. The implication of these findings is that civic skills (the sorts of skills required to deliberate) require *development* and, importantly, *practice* for one to become effective and proficient in their use.

In a similar vein, research about the development of expert skill by psychologist K. Anders Ericsson (1993, 2006) suggests that expertise is a function of repeated practice over time—but deliberate, goal-oriented practice that includes *feedback*. Of course, I am not suggesting that we make all citizens into deliberative experts (there is nowhere near enough time or resources for that), but the *approach* discussed by Ericsson offers both a theory and a roadmap for understanding why and how practice makes better, if not perfect. Improved skillsets allow for better deliberation, especially when one is training those skills in specific ways towards specific ends. As I have argued in previous chapters, encouraging certain *approaches* to deliberation and *practices* during deliberation, ought to help improve judgment and decision making.

Entrenching frequent and integrated deliberation would, I believe, encourage a move away from system 1/low-elaboration cognition and, specifically, would allow practitioners to focus on reducing participants' reliance on automaticity—i.e. autopilot thinking, which enables institutional shortcomings to thrive (e.g. the manipulation of citizen preferences by political parties and candidates/politicians, moneyed interests, etc.). This outcome would be especially likely if deliberative iterations included specific feedback related to epistemic goals. Frequent deliberation (a form of iteration) with those who think and reason differently allows for the cultivation of the sorts of deliberative goods I discussed in chapter five, such as a-rational receptivity, cognitive diversity, and targeted motivation, which is necessary to improve the degree to which individuals listen to and actively engage with/pay attention to their fellow deliberators. Allowing for more frequent and integrated deliberations also provides plenty of

opportunity for iteration, which should be useful for stabilizing preferences over time and building trust—each of which has a dual-use function.

Stable preferences and trust not only serve to generate better, more consistent, and more reliable judgments and decisions; they also further underwrite the very deliberative institutions that enable them in the first place, lending credibility to them and further legitimizing them as useful and appropriate democratic institutions available to supplement more traditional forms of politics. Indeed, if an approach to a deliberative system like that offered by Hendriks is taken—that of an “integrated system of public deliberation,” then by treating several locations within the system as deliberative spaces, citizens will have a plurality of recognized spaces for practicing deliberation and generating the goods I have mentioned here (Hendriks 2006; cf. a two-track approach in Habermas 1996). Hendriks argues that this approach would require that there is a celebration of the “*multiplicity* of deliberative venues” and that it would “foster *connections* between these venues” (Hendriks 2006: 499-500, emphasis in original). Such venues would include “parliaments, committee meetings, party rooms, stakeholder roundtables, expert committees, community fora, public seminars, church events, and so on” (Hendriks 2006: 499). Of course, we might not expect that each venue and instance of deliberation will yield much in terms of high-quality deliberation, or even a learning opportunity, but the hope is that such an approach will yield, on balance, a deliberative spirit (perhaps contingent on or moderated by approaches outside of these venues to generating citizen deliberative capacity). Within the system would be overlapping discursive spheres of formal, informal, and mixed discourse featuring deliberative and non-deliberative elements.

Now, within such a deliberative (or rather ‘discursive’) system, individuals would potentially have several, varied sites for engagement—for the *practice* of democratic engagement. Returning to the concept of practical wisdom—which I have argued is an approach to rationality—this sort of distribution of venues for practicing citizenship is essential for building capacities such as autonomy and rationality. If many of, perhaps even some of, the particular deliberative events within this system are structured in such a way as to encourage good judgments, then individuals would be given several opportunities to practice citizenship and transport their skills from one venue to another.

Perhaps most importantly of all, committing to a *deliberative democratic system* that enables and multiplies opportunities for substantive deliberative engagement over time helps address the institutional challenges to good judgment and decisions that I discussed in chapter six. By creating a robust (and sustainable) deliberative system, we are creating *institutions* that can produce democratic goods, enhance deliberation, and, in so doing, militate against the deleterious effects of other institutions that undermine good judgment and decision making by exploiting, deliberately or otherwise, the limits of our cognitive capacities.

### **Situated and balanced technocracy**

A deliberative system would allow for institutional changes—either universal or located in specific sorts of deliberative outlets or practices—that would balance expertise or technocratic approaches to decision making with more citizen-based approaches. Within any modern mass democratic system, there will be an extensive need for experts to manage the complexity that emerges from the massive scale, speed, and sophistication of contemporary social, political,

and economic issues. A deliberative system that includes several varied but interconnected sites of and approaches to political life and decision making allows for a balance, as we saw above, between expertise and technocracy on the one hand, and citizen-led action on the other.

As Parkinson (2006) argues, in order to achieve this balance, within deliberations moderators/facilitators should be trained to balance technocratic and citizen power; decisions made by citizens should be binding, or at least taken very seriously by senior decision makers (see Fuji Johnson 2015); and, finally, the initiation of deliberation should also occur at sites outside of the purview of public managers (151). Because variables such as complexity, available resources including time, the speed at which a decision must be taken, and who is included as being affected by an outcome—and thus in need of inclusion in deliberations that bring about an outcome—different types and scales of deliberation will be required. And just as much, different combinations of deliberative and non-deliberative mechanisms will be required. Within a deliberative system, not only can we better understand how to balance expertise and technocracy *within* a deliberation, but also we can also better understand how to achieve such a balance *across* sites of deliberation and non-deliberation (since there will be an implicit, or perhaps explicit, dialogue between the sites). This balance could contribute to better epistemic deliberation by helping ensure that the capacities of ordinary citizens are not atrophied while also ensuring that the necessary and significant role of experts and technocrats remains intact *on aggregate and in general*, despite some variance in their levels of engagement depending on the particular context surrounding the issue at hand.

Balancing the input of experts and citizens allows the public to have access to the information they need to make good judgments and decisions without presupposing that the opinions of experts have a pre-political value that trumps the will of citizens or their elected representatives. Expertise can contribute to facilitation by enabling moderators/facilitators to better engage participants in deliberation. And if the background noise that contributes to cognitive distortion is minimized through other techniques—the sorts discussed in chapter five—those who deliberate will be better suited to listen to expert advice and incorporate it into their judgments and decisions. As I argued in previous chapters, the *primary* nature of the challenge of distorted judgment and decision making under deliberative (or other) conditions is not that citizens *lack any capacity* to think, judge, and choose wisely; rather, it is that the sorts of contexts, incentive/disincentive structures, rules, and institutions that decision makers typically find themselves within tend to exacerbate cognitive vulnerabilities and encourage distorted or biased thinking. That is why one of the major themes of this dissertation has been that changing structures is essential; that way, functions (i.e. practices, and therefore outcomes) can change, and work can be done to cultivate and enhance citizen capacity for autonomy and rationality.

The question then becomes: Where do we draw the line with technocracy and expertise in democratic deliberation? How much is too much? At what do the managers and experts cease to be servants of the public and instead become their (de facto) managers? While a continually-negotiated balance will likely be required for each deliberation, there are two guiding principles, noted about in reference to Christiano, that ought to be followed: first citizens should remain, broadly speaking, in the control of the agenda; and second, (elected)

policy-makers ought to be ultimately responsible and accountable for taking formal decisions. To these guiding principles, and in pursuit of high-quality deliberations, I add that it ought to be *those who deliberate* who get to decide when to call on experts and they ought to have the right to request other experts as needed during a given deliberation outside of the basic provision of experts/introductions set up by those hosting the deliberation. Organizers and facilitators should remain in charge of recommending, preparing, and providing experts for consultations during a deliberation, but it should be the deliberative collective who maintain control of the who and when as deliberation unfolds (especially as deliberators become more familiar and comfortable with the issue(s) they are deliberating). Finally, the idea that experts are provided for *consultation* and not *direction* is important and must be maintained during a deliberation; with the former, it is participants who lead the experts rather than experts who lead participants. Respecting this principle helps ensure that no technocracy-by-stealth is introduced into deliberation (or that it is, at the very least, limited).

Institutionally, balancing technocratic/expert engagement against other intervention is important—especially if you can sort the experts and technocrats from other interveners who may try to claim the status while harbouring some ulterior motive (e.g. the television “expert” who is really just a shill for a political party). This balance is particularly important both inside and outside of deliberation since the sorts of sensibilities, biases, opinions, and preferences that individuals have *before and after* deliberation will be affected by the flow of information that surrounds (i.e. precedes and follows) a deliberative event. For instance, electoral campaigns are full of interested parties flooding the airwaves, sidewalks, and the Internet, each advancing an agenda. The same is true, though perhaps to a lesser degree, outside of an

electoral period. To the extent that *genuine* technocratic/expert intervention can be institutionalized, amplified, and made accessible both within and outside of deliberation, citizens will be better prepared to counter messages and information shared by those whose interest is persuading individuals of the value of some proposition or communicating some information without concern for the ideal democratic deliberative requirements of reason-giving, reciprocity, revisability, publicity, autonomy, and rationality.

While it is perfectly reasonable to expect such strategic political activity to occur in a liberal democracy marked by several approaches to practicing democratic politics, it is also important to counter the potentially negative effects and influences of strategic/tactical politics aimed at producing pre-determined outcomes favourable to some group or interest that may or may not be good for the public at large. If politics is competitive, then democratic deliberation itself serves as a competitive, alternative, and yet none-the-less-complementary approach to engaging politically. Experts and technocrats are a key part of making this alternative appealing. As I have noted, balancing the role and influence of technocrats/experts outside of deliberation will serve deliberative politics both internally (through the direct engagement of these individuals as witnesses/educators) and externally (through the indirect engagement of these individuals as public communicators of ideas and agenda setters). That said, given that a) there is no such thing as an unbiased interlocutor (only a more or less biased one) and b) that there is no pre-political technocratic or expert value when it comes to deliberation, final decisions, regardless of outside intervention, ought to be left to the citizens or their representatives. However, the better served they are by having access to good information *before judging or deciding*, the better for the public and for democracy.

## Stability and disruption

A properly designed deliberative system would provide enough stability so that, if properly set up, the rules of the game remain predictable and reliable, reducing the long-term cost of becoming involved (i.e. learning the rules, navigating the system); this approach also improves the long-term publicity and familiarity of the deliberations themselves (Parkinson 2006: 158). And if the system allows for its rules to go through periodical review, a balance can be struck between stability and the sort of adaptability required of any just system within a socially and politically dynamic space (Goodin 1996: 40-41). Overall, properly balanced, a deliberative system should be familiar and accessible, such that the threats to ontological security and meaning maintenance, discussed earlier, are reduced, and the epistemic value of deliberative judgments are not compromised through the sorts of cognitive distortion that are generated or heightened under such conditions. This should, again, enable the long-term stabilization of preferences and trust building.

In a well-designed and functioning deliberative system, we might also imagine a macro-balance between the stability function and the sorts of disruptions required to kick participants off of “auto-pilot” and into higher or more critical, reflective, and attentive modes of cognition—such as those discussed in chapter five and in the two sections just above. Parkinson (2006: table, 169) lists seven kinds of elements of a deliberative system: activist networks, experts, the bureaucracy, micro techniques (e.g. deliberative polls and juries), the media, an elected assembly, and direct techniques (e.g. referenda and petitions)—an expanded list from the one detailing elements of a deliberative system that we saw at the beginning of this chapter. The rules embedded within specific elements of this system can be developed in such

a way as to meet the requirements of stability and revisability mentioned just above, but, when combined, a kind of critical tension should emerge. This tension, if properly harnessed by the system, should help improve the epistemic value of deliberation by enhancing the techniques and designs *within individual deliberations* discussed in chapter five, since individuals will have to contend with multiple sources of information and sites of political engagement. Why? Because while some *predictability of procedure* is important for generating stronger cognitive capacities and better practices, some *friction within deliberation itself* is necessary to prevent individuals from relying on off-the-shelf heuristics or habitual, non-reflective, or uncritical patterns of cognition embedded in psychological phenomena including motivated reasoning, system justification, automaticity, and social intuitionism.

Institutionally, a deliberative system would allow multiple points of contestation, such that internal institutions would be pluralized and less subject to hijacking by particular interests. While a well-functioning system would be protected from such external challenges—such as electoral politics, partisan intervention, and some of the more egregious perversions of hegemony and ideology—internally, well-designed deliberative events would allow for the exploration of issues, questions, and preferences related to elements found in and emerging from these other institutions. For instance, the question of whether a national senate should be a partisan body is an issue that, outside of deliberation, provokes a debate that is almost surely ripe to be full of partisan and entrenched interests. However, *inside a deliberative system* there would be spaces where a different sort of conversation could occur—one that might be more honest, productive, and constructive, and which could feed back into the broader discussion outside of particular deliberative events and the deliberative system as a whole. Once again,

this approach would serve both to mitigate the most deleterious effects of individual cognitive biases and distortions *and* institutional perversions. So, a deliberation that produces a recommendation for a particular sort of senate might serve as a heuristic for the broader public; such a recommendation might be seen as the product of a disinterested or semi-disinterested public exploration of technical information and normative considerations. In a sense, the product of a deliberative event, and more broadly the products of deliberative institutions, can be seen as producing recommendations from an honest broker or brokers. While such an output might not be externally binding, it would at least be instructive and likely reliable.

### **Experiments in deliberation**

As I mentioned in chapter five, and as I will revisit in my conclusion, theories about deliberation and cognition *must be tested against current and changing political and social realities* of whatever jurisdiction to which they are applied. What we know about what works and what does not when it comes to producing better epistemic outcomes through deliberation is currently limited and almost certainly subject to change. A deliberative system allows for the scale and scope required to experiment with different approaches to deliberation under a variety of conditions. Moreover, given that building a deliberative system is likely to be a long-term project, experts and researchers will be able to compare changes in deliberation across time as new or altered practices and designs are adopted. The process of generating better deliberation is exactly that: *a process*; and, indeed, it is a dialectical one. Practicing deliberation within a deliberative system allows for a bolder dialectical process and for the kinds of

experiments that will be required if we are to discover the best approaches to generating epistemically good judgment and decision making in deliberation. A deliberative system allows for experimentation at both the level of specific deliberative events, with a focus on how the design of such events affect the thinking and judgments of individuals and the decisions of the collective, and the level of deliberative institutions over time. This approach allows practitioners and researchers to, over time, identify and address the structural strengths and weaknesses of approaches to both specific deliberations and deliberative practices more generally.

Of particular note and importance is the fact that since, as I have argued, the sorts of cognitive biases and distortions that I have addressed in this dissertation—the sorts that undermine autonomy, rationality, and thus judgment and decision making—are in part enabled by institutions that exacerbate existing human cognitive shortcomings. By instituting a deliberative system as I have discussed here, with specific reference to Parkinson et al., counter-institutions can be entrenched to provide a systematic, persistent response to challenges raised by, among others, the psychological phenomena of automaticity, social intuitionism, system justification, and motivated reasoning: namely, system 1/low elaboration cognition gone awry. To counter the deleterious effects—when it comes to cognition, autonomy, and rationality—of these phenomena and the institutions that tend to generate them, including campaigns and elections, electoral systems, the media, political parties and partisanship, hegemony and ideology, one-off deliberative events will not be enough. Only a deliberative system, properly designed, can begin to do and to entrench the sort of work that needs to be done to counter these challenges.

Moreover, within a deliberative system the conceptual approaches to addressing these challenges that I have mentioned in this dissertation—a-rational receptivity, cognitive diversity, iteration, targeted motivation, and wise facilitation—can be themselves entrenched as part of the design. They can be tried, tested, and revised and needed. Their status and permanent or semi-permanent elements of engagement may also stand to generate new norms both within deliberations and outside of them, ideally norms commensurate with the normative goods prescribed by deliberative theory in particular and democratic governance in general. But such a move, of course, takes time, effort, study, and both trial and error. A deliberative system offers the environment necessary for all of this to occur.

## **Conclusion and summary**

In this chapter I have looked at different approaches to deliberative systems: what functions such a system would play, which elements it would include, and the kinds of deliberative venues that it might require. I outlined the specific roles of experts and citizens—and how they relate—in such a system, arguing that each plays a critical and interrelated role. I have also suggested that while we have only limited empirical data to examine the effects such an institution would have, we none the less have good reason to expect that a deliberative system, properly designed, would offer several epistemic benefits through enhancing the *practice* of deliberation and improving cognition as it relates to generating judgments on political matters and decisions related to those matters. I have also examined how a deliberative system would address the specific cognitive distortions and biases discussed in previous chapters, and how such a system would serve as a corrective to the deleterious effects of certain democratic

institutions in which these distortions and biases tend to become embedded. Finally, I discussed how a deliberative system would allow for new norms, approaches to judgment and decision making, and experiments in deliberation to emerge and become entrenched in a broader democratic system.

## Chapter 8: Practicing democratic deliberation

This dissertation has been about political judgment and decision making. It has been about the conditions under which good judgments and decisions are made by individuals who deliberate, and the conditions under which they are not. Specifically, I have asked whether individuals can make autonomous, rational judgments and whether deliberative democracy can serve as a theory of how to generate epistemically good judgments and decisions. I have argued that for theories of deliberative democracy, as they have been developed to date, to deliver the goods they promise, we require participants with fairly robust and developed capacities of rationality and autonomy (especially autonomy). I have also argued that there are reasons to be sceptical as to whether individuals who deliberate tend to exhibit these high levels of these capacities. I have suggested that there is a significant gap between what we tend to be capable of (nature)—whether by nature or nurture or both—and what is expected of us in theories of deliberation (culture). I have argued that this gap must be narrowed if theories of deliberation are to be fully developed and made to yield practices that can be institutionalized in a deliberative system that would be a part of the broader democratic culture.

I have shown how familiar political institutions outside of deliberation exploit our cognitive shortcomings and add to the difficulty of the reformist deliberative democratic agenda, and how building deliberative systems might serve as a check on or even a corrective to these institutions. In doing so, I have outlined why, for the question of our ability to make judgments and decisions that reflect our pursuit of survival and well-being, we must find ways of overcoming, or at least of mitigating, the effects of certain kinds of cognitive distortions and the flaws in institutions that allow them to be exploited. I have offered some tentative concepts

about how specifically this might be done at both the level of personal practice and institutional design.

This dissertation has also been about human potential. As a fundamentally normative project, the argument I have made about the gap between what we expect from people and what such people tend to be able to deliver has not been about abandoning our most ambitious hopes for democratic participation. Rather, it has been about finding ways to further empower individuals and to increase the number of those who can and want to deliberate—and who can deliberate well within the context of a given deliberative event or system. So, this dissertation has also, implicitly, asked the question: Should we make a normative commitment to improving the quality and extent of citizen participation in law and policy formation through deliberation? Of course, the answer, as far as I am concerned, is yes, and I have suggested some ways to bring about this change.

### **Why bother with epistemically good democratic deliberation?**

Why should we care? As I argued, the importance of democratic deliberation to self-determination is clear. Democracy is largely about collective self-determination, and this includes providing space for individual self-determination. To the extent that autonomy and rationality are undermined by a-rational cognitive phenomena occurring outside the awareness and control of individuals, these endeavours of self-determination are at risk. By enabling capacities of rationality and autonomy to flourish, individuals and collectivities will be better placed to decide what they want to do, why they want to do it, and how they should go about translating their will into concrete outcomes.

But self-determination can be *practiced* in better or worse ways. I have also argued that democratic deliberation *already* tends to produce better outcomes in many instances than alternative approaches to democratic decision making (e.g. aggregative democracy). But deliberation faces threats from the a-rational phenomena I have discussed. Such threats apply most notably to the *epistemic* defense of deliberation since they risk undermining the validity upon which deliberation relies for public judgments and decisions. When I say that self-determination can be practiced in better or worse ways, I mean that the quality of decisions that are reached, which enable self-determination, can vary according to the *process* used to bring them about. So, another reason why we should care about epistemically good democratic deliberation is that not only does such a practice allow for self-determination, it allows for *better self-determination*. By this, I mean that it allows for groups of people to better establish valid, just, and legitimate outcomes through reason giving. However, to maximize the potential of democratic deliberation to yield high-quality decisions, changes must be considered at the levels of individual practice, deliberative design, and deliberative system building.

## **Implications and further research**

Going forward in this field of research, I think we have good reason to be cautiously optimistic about what we might be able to do to generate better politics: that is, a politics that is marked by more common and deeper individual engagement, more diverse and fair representation of both policies and representatives, and more legitimate and better-understood policies and laws. New and ongoing research in the fields of cognitive neuroscience, behavioural economics, social psychology, political psychology, and political theory is helping to reveal further how and

why we behave the way we do. Moreover, such research is also helping us understand the specific ways in which the environment, including the institutions that have developed over time—deliberately, accidentally, or a mix of both—to assist us in ordering and navigating this environment, intersects with cognition to produce behaviour. As we come to understand more and more how such things operate, we can come to understand how we should attempt to order things so that we can *produce better judgments and decisions*; and while it is outside of the scope of this dissertation, I would argue that these better judgments and decisions might produce better outcomes for individuals who are currently effectively disenfranchised through certain political processes that undermine the goals, norms, and requirements of a democratic system—so the epistemic goods of deliberation may, indeed, be significant for other social and political reasons.

It is important that the discussion about this research that emerges is one about *better or worse* outcomes. Ongoing, open, and revisable work must be done to define what counts as better or worse and under what conditions. It is just as important that researchers and the public understand that there is no perfect outcome, since, on the one hand, it seems unlikely that we will ever fully overcome our cognitive limitations, and, on the other hand, it is just as unlikely that we will ever agree on what counts as perfect. Still, on some reasonably objective measures—should those bound up in a given political body choose them as normatively desirable ends—we can assess attempts at improving democracy by asking whether the practices, institutions, laws, and rules of a democracy enhance or detract from democratic self-governance.

For instance, we can ask: Are more or fewer individuals able to substantively engage in effective deliberation? Is participation marked by more or less autonomous and rational engagement (perhaps according to the standards of evaluation laid out in chapter six)? Are individuals who wish to engage given adequate resources to enable them to do so? Are policies more or less reflective of the stated preferences of informed citizens? In sum: Are the tools of democratic self-governance reflective and effective?

This project has engaged measures of deeper democracy in a few ways. The most obvious is that some of the conceptual evaluation tools that have been discussed raise questions about the *quality* of participation—that is, how engaged are participants, how extensive is their political knowledge, are they able to translate facts about the world, their values, and their preferences into constructive dialogue towards some end. Others are about *accessibility* and *receptivity*—that is, are deliberative events and systems open to participation. The former class is obviously linked insofar as the questions are about whether participation and especially deliberation, becomes better or worse after we pay attention to reforming and improving democratic institutions and the attendant practices of the individuals who engage with them. The latter are more circuitously related.

The implicit expectation that links this second set of evaluative questions to the goals of this project is broadly rooted in critical theory tradition surrounding the ideas of freedom and self-government. More specifically, one of the underlying normative drivers of this project has been the idea that individuals who are better able to develop and employ the capacities of autonomy and reason—in both the basic rational choice and practical wisdom senses I have been using it—are more likely to have their preferences reflected in the social and political

makeup of the worlds in which they live. Just as important is the idea that if we can even out the balance of such capacities, we can better reach a state of social and political equilibrium, challenging the current state of affairs in which the preferences of some are strongly considered and more likely to be adopted than those of others.

To the structuralist critique or to those who would invoke the overwhelming forces of liberalism and capitalism as meta-institutions that shape so much of our behaviour, and therefore make the generation of capacities of autonomy and rationality secondary endeavours, I can only say at this point that developing tools must be done prior to building something—that is, any project of resistance or reform will require capacities for practicing resistance or bringing about that reform. If peaceful reform is what certain progressive individuals and movements are after, then sophisticated deliberative politics—enabled in part by heightened democratic capacities—is a good place to start. Again, while deliberative events and systems are *a part of* a broader political system that also involves parliaments and protests, civil disobedience and letter writing campaigns, and other sites of engagement, deliberation can become a very important site of political engagement.

Moreover, this project does not assume *what* individuals would want to choose as preferences or *how* they would want their social and political world to look. It only assumes that it is likely they have the *tools to develop preferences, to have good reasons for those preferences that can be shared with others, and to see them adopted—or at least treated and considered seriously*. Generating individual capacities to better able to develop and communicate preferences and reasons for those preferences, through deliberation, harbours social justice potential bound up with the goods of democratic self-government. Ultimately, the

ability to have good reasons for and against preferences, to publicly share those with others, and to have them feed into the public sphere—and perhaps adopted as law or policy—is the essence of such self-government. To the extent that this process is undermined, including the capacities for rationality and autonomy that underwrite this process, democratic self-government itself is undermined.

## **Next steps**

It is now well known that individuals are not the Bayesian calculators and updaters or the end-means maximizers of certain economic theoretical lore. Such rationalist approaches set up the question of how we *might* or *do occasionally* think about things, but data from social and political psychology, from cognitive neuroscience, from sociology, and from other behavioural and social sciences have given evidence to demand we revisit our understanding of how we think and why we behave in the ways we do. Even when such models are meant to be *aspirational*, rather than *descriptive*, something is lost. As I have noted, practical wisdom approaches to judgment and decision making are valuable, and capture something productive, democratic, and human that are lost when strict rationalist models are privileged or relied upon exclusively in the study of self-government. However, going too far the other way—into the realm of primarily affective judgment decision making—is no better.

Nonetheless, within the understanding of humans, as fundamentally imbued with affect, cognitively bounded, and highly externally conditioned beings is the potential for us to restructure our institutions to best suit our collective capacities. Said another way, the more we know about how affective processes influence our political and social decision making, the

better we can adjust our systems and the institutions that comprise them, the more we can adjust our personal practices and ways of thinking in ways that encourage better judgments and decisions that we can mobilize towards goals of equality and dignity (Damasio 1994).

This project provides the foundation for the future development and testing of hypotheses related to deliberative practices and systems. Any hypotheses that emerge from the concepts I have developed or added to will need to be tested and revised in an ongoing, revisable critical dialogue about how we behave under various conditions, how we can do better, and, of course, what counts as better and why. Broader emergent questions include how and why democratic institutions must be reimagined and then reconfigured according to deliberative strategies. These steps towards hypotheses—provisional and made with the intent for use by future experiments, which would consider emerging findings from behavioural sciences—will also have to remain flexible and those who develop them should be prepared to continue to incorporate new findings from the behavioural sciences.

I have offered a critique of deliberation as a political theory of epistemic judgment and decision making, but I have also presented concepts from other theorists that, developed together here, point to institutional reform. Taking these and building upon them, refining them, and developing ways of testing them against the realities deliberation will be essential to making this theory work. The critique I have made of our capacity to generate autonomous, rational judgments and decisions in deliberative contexts is not a white flag; it is a starter's pistol. The concepts I have offered about each other are the first step in my project of working to close the gap between what we expect of deliberators and what, as their most empowered

selves, they are encouraged to deliver. This project draws from and builds upon a long tradition of behavioural critiques and studies and both theoretical and practical work on deliberation.

## **The place of deliberation in our political lives**

The practice of deliberative democracy is not a panacea for all our concerns about where citizens fit into their democracy, and we should not try to make it one. The challenges of life in contemporary democracies marked by diversity, persistent disagreement, complexity, and the speed of contemporary societies and the political lives enabled within them. As I have noted, reforms to representative democracy, including electoral reform, are required for citizens to see themselves as more accurately and proportionately reflected in formal systems of power. Reforms are also needed to deliver more substantive opportunities for citizens to practice democracy, to understand how it works, and to influence policy outcomes to make them more representative of the whole and more responsive to their needs. Protests, civil disobedience, petitions, and other forms of activism will remain important, and they may indeed be enhanced through more rigorous understandings of political behaviour and social cognition.

Nonetheless, when an important or difficult policy decision must be made—or one that morally requires substantive input from the people—some degree of popular deliberation is a useful approach for all levels of government, whether municipal, provincial or state, and national. As an approach to both preference formation and policy generation, deliberation, carefully designed with inclusive principles in mind, allows individuals the time and other resources required to maximize the chance that a good decision is made. Equally as important as the deliberation itself is what is generated through the process beyond the decision itself. As

was the case in the 2004 Citizens' Assembly on Electoral Reform carried out in British Columbia and as is the case in the state of Oregon Citizens Initiatives Review, the outcomes of deliberation may become useful heuristics for the rest of the population (and even *more useful*, I would argue, if those deliberations are carried out in ways that account for the critiques I have raised in this dissertation). How? At the very least, citizens will have an issue-educated peer-body they can consult in search of guidance on an issue (or emulate). In this sense, deliberations become, to paraphrase Arendt, islands of heuristic stability in an ocean of dubious information and cues.

Admittedly, deliberative democracy, as I have characterized it in these pages, is occasionally exclusionary (though it is often much more inclusive than other approaches to democratic decision making). As I have mentioned, not everyone deliberates according to these models and forms of communication, and some will not wish to, perhaps based on historical and persistent distrust of formal systems. To the former admission, I can only reply by noting that deliberation is *one approach to democratic participation* that should exist among many. Some reform, such as voter support and access, will be more inclusive, and to the extent they are, they will be superior approaches for those who are unable or otherwise disinclined to deliberate. Improving opportunities for, and the quality of, deliberative participation is still a valuable and important endeavour insofar as it fosters the potential to ameliorate both a disengaged and potentially disillusioned public and the corrupt political practice of catering to affective responses that are yet below the surface when it comes to some political activity, such as voting or indicating a policy preference.

## **Deliberation, decency, and human capacity**

One of the concerns addressed in this dissertation has been with the challenge of how we can leverage our understanding of human cognition and behaviour to bring about better political judgments and decisions through enhanced capacities for autonomy and rationality practiced through deliberative democracy—if we can at all. As mentioned above, I think we can. This project is thus situated within the tradition of critical political theory and aimed at thinking about problems of human freedom; it is about using the power of reason—not necessarily Bayesian rationality or ends-means utility maximizing rationality—to bring about more representative, fair, and just political outcomes. Underlying this project is the assumption that most human beings are fundamentally capable of rational, autonomous deliberation (and debate), but nonetheless susceptible to carrying out far-less-than-such behavior—especially when the institutions in which they live encourage poor thinking. I see much of this behaviour as being the outcome of ways of thinking that could, in theory, be heavily mitigated through further improving our democratic institutions and the personal practices that surround and shape them.

## Bibliography

Albright, Linda, David A. Kenny, and Thomas E. Malloy. "Consensus in Personality in Judgments at Zero Acquaintance." *Journal of Personality and Social Psychology*, vol.55, issue 3 (September 1988): 387-395.

Althaus, Scott L. "Information Effects in Collective Preferences." *The American Political Science Review*, vol.92, no.3 (September 1998): 545-558.

Ambady, Nalini and Robert Rosenthal. "Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences: A Meta-Analysis." *Psychological Bulletin*, vol.111, no.2 (March 1992): 256-274.

Ang, Ien. *Living Room Wars: Rethinking Media Audiences for a Postmodern World*. New York: Routledge, 1996.

Aristotle. *The Nicomachean Ethics*. Translated by J.A.K. Thomson. New York: Penguin, 2004 (circa 350 BCE).

Ballew, Charles C. II and Alexander Todorov. "Predicting Political Elections from Rapid and Unreflective Face Judgments." *Proceedings of the National Academy of Sciences*, vol.104, no.45 (November 2007): 17948-17953.

Ball, Terence and Richard Dagger. *Political Ideologies and the Democratic Ideal*. Pearson, 2010.

Bargh, John. "Introduction." In *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes*, edited by John Bargh. New York: Psychology Press, 2007.

Bargh, John and Kay L. Schwader, Sarah E. Hailey, Rebecca L. Dyer, and Erica J. Boothby. "Automaticity in Social-Cognitive Processes." *Trends in Cognitive Sciences*, vol.16, no.12 (December 2012): 593-605.

Bargh, John and Ezequiel Morsella. "The Unconscious Mind." *Perspectives on Psychological Science*, vol.3, no.1 (2008): 73-79.

Bargh, John and Tanya L. Chartrand. "The Unbearable Automaticity of Being." *American Psychologist*, vol.54, issue 7 (July 1999): 462-479.

Barsade, Sigal G. "The Ripple Effect: Emotional Contagion and its Influence on Group Behavior." *Administrative Science Quarterly*, vol.47, no.4, (December 2002): 644-675.

Bartels, Larry M. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior*, vol.24, no.2 (June 2002): 117-150.

Batalha, Luisa, Nazar Akrami, and Bo Ekehammar. "Outgroup Favoritism: The Role of Power Perception, Gender, and Conservatism." *Current Research in Social Psychology*, vol. 13, no.4 (2007): 39-49.

Bauer, Paul C., Pablo Barberá, Kathrin Ackermann, and Aaron Venetz. "Is the Left-Right Scale a Valid Measure of Ideology? Individual-Level Variations in Associations with 'Left' and 'Right' and Left-Right Self-Placement." *Political Behavior* (2015): 1-31.

Beiner, Ronald and Jennifer Nedelsky. "Introduction." In *Judgment, Imagination, and Politics*, edited by Ronald Beiner and Jennifer Nedelsky. Maryland: Rowman and Littlefield, 2001.

Belanger, Eric and Laura Stephenson. "Parties and Partisans: The Influence of Ideology and Brokerage on the Durability of Partisanship in Canada." In *Voting Behaviour in Canada*, edited by Laura Stephenson and Cameron D. Anderson. Vancouver: University of British Columbia Press, 2010.

Benhabib, Seyla. "Toward a Deliberative Model of Democratic Legitimacy." In *Democracy and Difference*, edited by Seyla Benhabib. Princeton: Princeton University Press, 1996, 67-94.

———. *Situating the Self: Gender, Community, and Postmodernism in Contemporary Ethics*. New York: Routledge, 1992.

Berelson, Bernard, Paul Lazarsfeld, and William McPhee. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press, 1954.

Berofsky, Bernard. *Liberation from Self: A Theory of Personal Autonomy*. New York: Cambridge University Press, 1995.

Bettencourt, Ann B. and Nancy Dorr. "Cooperative Interaction and Intergroup Bias: Effects of Numerical Representation and Cross-Cut Role Assignment." *Personality and Social Psychology Bulletin*, vol.24 (1998): 1276-1293.

Bizer, George Y., Zakary L. Tormala, Derek D. Rucker, and Richard E. Petty. "Memory-based Versus On-line Processing: Implications for Attitude Strength." *Journal of Experimental Social Psychology*, vol. 32 (2006): 646-653.

Black, Laura W., Stephanie Burkhalter, John Gastil, and Jennifer Stronmer-Galley. "Methods for Analyzing and Measuring Group Deliberation." In *The Sourcebook for Political Communication and Research: Methods, Measures, and Analytic Techniques*, edited by Erick P. Budy and R. Lance Holbert. New York: Routledge, 2010: 323-345.

Blais, Andre, Maxime Heroux-Legault, Laura Stephenson, William Cross, and Elisabeth Gidengil. "Assessing the Psychological and Mechanical Impact of Electoral Rules: A Quasi-Experiment." *Electoral Studies*, vol.31 (2012): 829-837.

Bohman, James. "The Coming Age of Deliberative Democracy." *The Journal of Political Philosophy*, vol.6, no.4 (1998): 400-425.

———. "Deliberative Democracy and the Epistemic Benefits of Diversity." *Episteme: A Journal of Social Epistemology*, vol.3, issue 3 (2006): 171-191.

Brady, Henry E., Sidney Verba, and Kay Lehman Schlozman. "Beyond SES: A Resource Model of Political Participation." *American Political Science Review*, vol.89, no.2 (June 1995) 271:294.

Brandom, Robert. *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, Mass.: Harvard University Press, 1994.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. *The American Voter*. New York: Wiley and Sons, 1960.

Cameron, Maxwell. *Between Rules and Practice: Why We Need Practical Wisdom in Politics*. Unpublished Manuscript, 2014

Chartrand, Tanya and John Bargh. "The Chameleon Effect: The Perception-Behavior Link and Social Interaction." *Journal of Personality and Social Psychology*, vol.76, no.6 (June 1999): 893-910.

Chong, Dennis and James N. Druckman. "Framing Public Opinion in Competitive Democracies." *American Political Science Review*, vol.101, issue 4 (2007): 103-126.

Christiano, Thomas. "Deliberation Among Experts and Citizens." In *Deliberative Systems*, edited by John Parkinson and Jane Mansbridge. Cambridge: Cambridge University Press, 2012: 27-51.

Christman, John. "Autonomy and Personal History." *Canadian Journal of Philosophy*, vol.21, no.1 (March 1991): 1-24.

Cohen, Geoffrey L. "Party Over Policy: The Dominating Impact of Group Influence on Political Beliefs." *Journal of Personality and Social Psychology*, vol.85 (2003): 808-822.

Cohen, Joshua. "Deliberation and Democratic Legitimacy." In *Democracy*, edited by David Estlund. Massachusetts: Blackwell, 2002.

Converse, Philip E. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, edited by David E. Apter. New York: The Free Press of Glencoe, 1964.

Cutler, Fred. "The Simplest Shortcut of All: Sociodemographic Characteristics and Electoral Choice." *The Journal of Politics*, vol.64, no.2 (May 2002): 466-490.

Dahl, Robert A. *Democracy and its Critics*. New Haven: Yale University Press, 1989.

Damasio, Antonio. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Random House, 1994.

———. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. New York: Harcourt, 2003.

Dasgupta, Nilanjana. "Implicit Ingroup Favoritism, Outgroup Favoritism, and Their Behavioral Manifestations." *Social Justice Research*, vol. 14, no.2 (June 2004): 143-169.

De Dreu, Carsten. "When Too Little or Too Much Hurts: Evidence for a Curvilinear Relationship Between Task Conflict and Innovation in Teams." *Journal of Management*, vol. 32 (2006): 83-107.

Denis, Laura. "Kant and Hume on Morality." *The Stanford Encyclopedia of Philosophy*. (Winter 2014 edition). Edward N. Zalta, editor. URL: <http://plato.stanford.edu/archives/win2014/entries/kant-hume-morality>.

Denzau, Arthur T. and Douglass C. North. "Shared Mental Models: Ideologies and Institutions." *Kyklos*, vol.47, issue 1 (1994): 3-31.

De Vries, Raymond, Aimee E. Stanczyk, Kerry A. Ryan, and Scott Y.H. Kim. "A Framework for Assessing the Quality of Democratic Deliberation: Enhancing Deliberation as a Tool for Bioethics." *Journal of Empirical Research on Human Research Ethics*, vol.6, no.3 (September 2011): 3-17.

De Vries, Raymond, Aimee E. Stanczyk, Ian F. Wall, Rebecca Uhlmann, Laura J. Damschroder, and Scott Y.H. Kim. "Assessing the Quality of Democratic Deliberation: A Case Study of Public Deliberation on the Ethics of Surrogate Consent for Research." *Social Science and Medicine*, vol.70, issue 12 (June 2010): 1896-1903.

Dijksterhuis, AP, Tanya L. Chartrand, and Henk Aarts. "Effects of Priming and Perception on Social Behaviour and Goal Pursuit." In *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes*, edited by John Bargh. New York: Psychology Press, 2007.

Ditto, Peter H. and David F. Lopez. "Motivated Skepticism: Use of Differential Decision Criteria for Preferred and Nonpreferred Criteria." *Journal of Personality and Social Psychology*, vol.63, no.4 (1992): 568-584.

- Druckman, James N. "Political Preference Formation: Competition, Deliberation, and the (Ir)relevance of Framing Effects." *The American Political Science Review*, vol.98, no.4 (November 2004): 671-686.
- . "On the Limits of Framing Effects: Who Can Frame?" *The Journal of Politics*, vol.63, no.4 (November 2001): 1041-1066.
- . "The Politics of Motivation." *Critical Review: A Journal of Politics and Society*, vol.24, no.2 (2012): 199-216.
- Dryzek, John S. "Democratization as Deliberative Capacity Building." *Comparative Political Studies*, vol.42, no.11 (November 2009): 1379-1402.
- . *Discursive Democracy*. New York: Cambridge University Press, 1990.
- Dunning, David. "Motivated Cognition in Self and Social Thought." In *APA Handbook of Personality and Social Psychology: Volume 1 Attitudes and Social Cognition*, edited by Mario Mikulincer and Phillip R. Shaver. Washington, D.C.: American Psychological Association, 2015.
- Duverger, Maurice. *Les Partis Politiques*. Paris: Armand Colin, 1951.
- Dworkin, Gerald. *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press, 1988.
- Eagly, Alice H. and Shelly Chaiken. *The Psychology of Attitudes*. Orland: Harcourt Brace, 1993.
- Edelman, Gerald M. *Second Nature: Brain Science and Human Knowledge*. New Haven: Yale University Press, 2006.
- Edwards, Peter B., Richard Hindmarsh, Holly Mercer, Meghan Bond, Angela Rowland. "A Three-Stage Evaluation of a Deliberative Event on Climate Change and Transforming Energy." *Journal of Public Deliberation*, vol.4, issue 1 (2008): article 6.
- Elstub, Stephen. "A Double-Edged Sword: The Increasing Diversity of Deliberative Democracy." *Contemporary Politics*, vol.12, issue 3-4 (2006): 301-320.
- . *Towards a Deliberative and Associational Democracy*. Edinburgh: Edinburgh University Press, 2008.
- Ericsson, K. Anders. "The Influence of Experience and Deliberate Practice on the Development of Superior Expert Performance," in *The Cambridge Handbook of Expertise and Expert Performance*, edited by K. Anders Ericsson, Neil Charness, Paul J. Feltovich, and Robert R. Hoffman. Cambridge, U.K.: Cambridge University Press, 2006.

Ericsson, K. Anders, Ralf Th. Krampe, and Clemens Tesch-Romer. "The Role of Deliberate Practice in the Acquisition of Expert Performance," *Psychological Review*, vol. 100 (1993): 363-406.

Erikson, Robert S. and Kent L. Tedin. *American Public Opinion, Sixth Edition*. New York: Longman, 2003.

Estlund, David. *Democratic Authority: A Philosophical Framework*. Princeton: Princeton University Press, 2008.

———. "Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority" in *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman, James and William Rehg. Cambridge: MIT Press, 1997.

Evans, Geoffrey and Mark Pickup. "Reversing the Causal Arrow: The Political Conditioning of Economic Perceptions in the 2000-2004 U.S. Presidential Election Cycle." *The Journal of Politics*, vol.72, no.4 (October 2010): 1236-1251.

Fishkin, James. *The Dialogue of Justice: Toward ad Self-Reflective Society*. New Haven: Yale University Press, 1996.

Flyvbjerg, Bent. *Rationality and Power: Democracy in Practice*, trans. S. Sampson. Chicago: University of Chicago Press, 1998.

Frank, Thomas. *What's the Matter with Kansas: How Conservatives Won the Heart of America*. New York: Metropolitan Books, 2004.

Freeze, Melanie and Jacob M. Montgomery. "Static Stability and Evolving Constraint: Preference Stability and Ideological Structure in the Mass Public." *American Politics Research*, vol. 44, no. 3 (May 2016): 415-447.

Fuji-Johnson, Genevieve. *Democratic Illusion*. Toronto: University of Toronto Press, 2015.

Gastil, John and Kanan Sawyer. "When Process Matters: An Exploration of Different Approaches to Operationalizing Public Deliberation." Conference Paper, Annual Meeting of the Western States Communication Association. Albuquerque, New Mexico: February 2004.

Gazzaniga, Michael. *Who's in Charge? Free Will and the Science of the Brain*. New York: HarperCollins, 2011.

Gerber, Marlène. "Equal Partners in Dialogue? Participation Equality in a Transnational Deliberative Poll." *Political Studies*, vol.63, issue S1 (April 2015): 110-130.

Gerber, Marlène, André Bächtiger, Susumu Shikano, Simon Reber, and Samuel Rohr. "Deliberative Abilities and Deliberative Influence in a Transnational Deliberative Poll." *British Journal of Political Science* (September 2016): 1-26.

Giddens, Anthony. *The Constitution of Society: Outline of a Theory of Structuration*. Berkeley: University of California Press, 1984.

Gidengil, Elisabeth. "Beyond the Gender Gap." *Canadian Journal of Political Science*, vol.40, issue 4 (2007): 1-17.

Gilens, Martin and Benjamin I. Page. "Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens." *Perspectives on Politics*. Vol.12, No.3 (2014): 564-581.

Goodin, Robert. "Institutions and Their Design." In *Theory of Institutional Design*, edited by Robert Goodin. Cambridge: Cambridge University Press, 1996.

———. "Sequencing Deliberative Moments." *Acta Politica*, vol.40 (2005): 182-196.

Gramsci, Antonio. *Prison Notebooks*, edited and translated by J.A. Buttigieg. New York: Columbia University Press, 1992.

Green, Donald, Bradley Palmquist, and Eric Schickler. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven: Yale University Press, 2002.

Gutmann, Amy and Dennis Thompson. *Democracy and Disagreement*. Cambridge: Belknap Press, 1998.

Gutmann, Amy and Dennis Thompson. *Why Deliberative Democracy?* Princeton: Princeton University Press, 2004.

Habermas, Jürgen. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, trans. William Rehg. Cambridge: Polity Press, 1996.

———. *On the Pragmatics of Communication*, trans. Barbara Fultner. Cambridge, MA: MIT Press, 1998.

———. *The Theory of Communicative Action*, trans. Thomas McCarthy. Boston: Beacon Press, 1984.

Haidt, Jonathan. "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review*. Vol.108, Issue 4 (2001): 814-834.

Hart, Philip S. and Erik C. Nisbet. "Boomerang Effects in Science Communication: How Motivated Reasoning and Identity Cues Amplify Opinion Polarization About Climate Mitigation Policies." *Communication Research*. 39(6) (2012): 701-723.

Hausknecht, Douglas R. and Danny L. Moore. "Time Compression, Response Opportunity, and Persuasion." *Journal of Consumer Research*, vol.13, no.1 (June 1986): 85-99.

He, Baogang and Mark E. Warren. "Authoritarian Deliberation: The Deliberative Turn in Chinese Political Development." *Perspectives on Politics*. Vol.9, Issue 02 (2011): 269-289.

Held, David. *Democracy and the Global Order: From the Modern State to Cosmopolitan Governance*. Cambridge: Polity Press, 1995.

Hendriks, Carolyn M. "Integrated Deliberation: Reconciling Civil Society's Dual Role in Deliberative Democracy." *Political Studies*, vol. 54 (2006): 484-508.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. "The Weirdest People in the World?" *Behavioral and Brain Sciences*, no.33 (2010): 61-83.

Heath, Joseph. *Enlightenment 2.0*. Toronto: HarperCollins, 2014.

Heine, Steven J., Travis Proulx, and Kathleen D. Vohs. "The Meaning Maintenance Model: On the Coherence of Social Motivation." *Personality and Social Psychology Review*, vol.10, no.2 (2006): 88-110.

Hill, Seth J., James Lo, Lynn Vavreck, and John Zaller. "How Quickly We Forget: The Duration of Persuasion Effects from Mass Communication." *Political Communication*, vol.30 (2013): 521-745.

Hong, Lu and Scott Page. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences*, no.46, vol.101 (2004): 16385-16389.

Hume, David. *A Treatise of Human Nature, second edition*, edited by Lewis A. Selby-Bigge and Peter H. Nidditch. Oxford: Oxford University Press, 1975 [1748].

Hurka, Thomas. *Perfectionism*. Oxford: Oxford University Press, 1993.

Iyengar, Shanto and Donald R. Kinder. *News that Matters: Television and American Opinion*. Chicago: University of Chicago Press, 1987.

Iyengar, Shanto, Mark D. Peters, and Donald R. Kinder. "Experimental Demonstrations of the 'Not-So-Minimal' Consequences of Television News Programs." *American Political Science Review*, vol.76, no.4 (December 1982): 848-858.

Jacobs, Lawrence R., Fay Lomax Cook, and Michael X. Delli Carpini. *Talking Together: Public Deliberation and Political Participation in America*. Cambridge: Cambridge University Press, 2009.

Jennings, Kent M. "Ideological Thinking Among Mass Publics and Political Elites." *Public Opinion Quarterly*, vol.56, no.4 (1992): 419-441.

Johnston, Richard. "Polarized Pluralism in the Canadian Party System: Presidential Address to the Canadian Political Science Association, June 5, 2008." *Canadian Journal of Political Science*, vol.31, issue 4 (December 2008): 815-834.

Jost, John T. "The End of the End of Ideology." *American Psychologist*, vol.61, issue 7 (2006): 651-670.

Jost, John T., Jack Glaser, Arie W. Kruglanski, and Frank J. Sulloway. "Political Conservatism as Motivated Social Cognition." *Psychological Bulletin*, vol.129, no.3 (2003): 339-375.

Jost, John. T. and David M. Amodio. "Political Ideology as Motivated Social Cognition." *Motivation and Emotion*, vol.36, (2012): 55-64.

Jost, John T. and Mahzarin R. Banjali. "The Role of Stereotyping in System-Justification and the Production of False Consciousness." *British Journal of Social Psychology*, vol.33 (1994): 1-27.

Jost, John T, Mahzarin R. Banaji, and Brian A. Nosek. "A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo." *Political Psychology*, vol.25, no.6 (2004): 881-919.

Jost, John. T., Nosek, Brian. A., and Gosling, Samuel. D. "Ideology: Its resurgence in Social, Personality, and Political Psychology." *Perspectives on Psychological Science*, vol.3, no.2 (2008): 126-136.

Kahan, Dan. "Ideology, Motivated Reasoning, and Cognitive Reflection." *Judgment and Decision Making*, vol.8, no.4 (July 2013): 407-424.

Kahneman, Daniel. *Thinking, Fast and Slow*. Toronto: Doubleday Canada, 2011.

Kahneman, Daniel and Amos Tversky. "On the Psychology of Prediction." *Psychological Review*, vol.80, no.4 (July 1973).

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor and Jens Timmermann. Cambridge: Cambridge University Press, 2012.

Kaufmann, Walter. *The Portable Nietzsche*. New York: Penguin Books, 1977.

Kay, Aaron C., Danielle Gaucher, Jennifer M. Peach, Kristin Laurin, Justin Friesen, Mark P. Zanna, and Steven J. Spencer. "Inequality, Discrimination, and the Power of the Status Quo: Direct Evidence for a Motivation to See Things As the Way They Should Be." *Journal of Personality and Social Psychology*, vol.97, no.3 (2009): 421-434.

Kay, Aaron C., Maria C. Jimenez, and John T. Jost. "Sour Grapes, Sweet Lemons, and the Anticipatory Rationalization of the Status Quo." *Personality and Social Psychology Bulletin*, vol.28, issue 12 (December 2002): 1300-1312.

Kim, Young Mie and Kelly Garrett. "On-line and Memory-based: Revisiting the Relationship Between Candidate Evaluation Processing Models." *Political Behaviour*, vol.34 (2012): 345-368.

Knight, Jack. "Models, Interpretations and Theories: Constructing Explanations of Institutional Emergence and Change." In *Explaining Social Institutions*, edited by Jack Knight and Itai Sened. Michigan: University of Michigan Press, 1995.

Kuklinski, James H. and Paul J. Quirk. "Reconsidering the Rational Republic: Cognition, Heuristics, and Mass Opinion." In *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, edited by Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin. Cambridge: Cambridge University Press, 2000, 153-182.

Kunda, Ziva. "Motivation and Inference: Self-Serving Generation and Evaluation of Evidence." *Journal of Personality and Psychology*, vol.53 (1987): 636-647.

Landemore, H el ene. "Deliberation, Cognitive Diversity, and Democratic Inclusiveness: An Epistemic Argument for the Random Selection of Representatives." *Synth ese*, vol.190, issue 7 (2013): 1209-1231.

— — —. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton: Princeton University Press, 2013.

Landemore, H el ene and Hugo Mercier. "Talking it Out with Others vs. Deliberation Within and the Law of Group Polarization: Some Implications of the Argumentative Theory of Reasoning for Deliberative Democracy." *Analise Social*, vol.xlvii, issue 4 (2012): 910-934.

Lenz, Gabriel. *Follow the Leader? How Voters Respond to Politicians' Policies and Performance*. Chicago: The University of Chicago Press, 2012.

— — —. "Learning and Opinion Change, Not Priming: Reconsidering the Priming Hypothesis." *American Journal of Political Science*, vol.53, no.4 (October 2009): 821-837.

Levendusky, Matthew. *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. Chicago: The University of Chicago Press, 2009.

Lodge, Milton and Charles Taber. *The Rationalizing Voter*. Cambridge: Cambridge University Press, 2013.

———. “Three Steps Towards a Theory of Motivated Reasoning.” In *Elements of Reason*, edited by Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin, Cambridge: Cambridge University Press, 2000.

Lodge, Milton, Marco R. Steenbergen, and Shawn Brau. “The Responsive Voter: Campaign Information and the Dynamics of Candidate Evaluation.” *The American Political Science Review*, vol.89, no.2 (June 1995): 309-326.

Lodge, Milton, Kathleen M. McGraw, and Patrick Stroh. “An Impression-Driven Model of Candidate Evaluation.” *The American Political Science Review*, vol.83, no.2 (June 1989): 399-419.

MacKenize, Michael K. and Kieran O’Doherty. “Deliberating Future Issues: Minipublics and Salmon Genetics.” *Journal of Public Deliberation*, vol.7, issue 1 (2011): article 5.

Mackenzie, Scott B. and Richard A. Spreng. "How Does Motivation Moderate the Impact of Central and Peripheral Processing on Brand Attitudes and Intentions?" *Journal of Consumer Research*, vol.18, no.4 (March 1992): 519-529.

Mansbridge, Jane, James Bohman, Simone Chambers, Thomas Christiano, Archon Fung, John Parkinson, Dennis F. Thompson, and Mark E. Warren. “A Systemic Approach to Deliberative Democracy.” In *Deliberative Systems*, edited by John Parkinson and Jane Mansbridge. Cambridge: Cambridge University Press, 2012: 1-26.

Mansbridge, Jane, Janette Hartz-Karp, Matthew Amengual, and John Gastil. “Norms of Deliberation: An Inductive Study.” *Journal of Public Deliberation*, vol.2, issue 1 (2006).

Mansbridge, Jane and Mark E. Warren. “Deliberative Negotiation.” In *Negotiating Agreement in Politics: American Political Science Association Task Force Report*, edited by Jane Mansbridge and Cathie Jo Martin. Washington: American Political Science Association, 2013, 86-120.

March, James G. and Herbert Simon. *Organizations*. New York: Wiley and Sons, 1958.

Martorana, Paul V., Adam D. Galinsky, and Hayagreeva Rao. “From System Justification to System Condemnation: Antecedents of Attempts to Change Power Hierarchies.” *Status and Groups: Research on Managing Groups and Teams*, vol.7, (2005): 285-315.

McAfee, Noelle. *Habermas, Kristeva, and Citizenship*. Ithaca: Cornell University Press, 2000.

Mello, Abby L. and Joan R. Rentsch. "Cognitive Diversity in Teams: A Multidisciplinary Review." *Small Group Research*, vol.46, no.6 (2015) 623:658.

Mendelberg, Tali. *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton: Princeton University Press, 2001.

Mercier, Hugo and H el ene Landemore. "Reasoning is for Arguing: Understanding the Successes and Failures of Deliberation." *Political Psychology*, vol.33, issue 2 (April 2012): 243-258.

Mill, John Stuart. *On Liberty*, edited by Edward Alexander. Peterborough: Broadview Press, 1999.

Moore, Alfred. "Following from the Front: Theorizing Deliberative Facilitation." *Critical Policy Studies*, vol.6, issue 2 (2012): 146-162.

Moore, George Edward. "On Truth and Falsity." In *Selected Writings*, edited by Thomas Baldwin. London and New York: Routledge, 1993.

Morris, William E. and Charlotte R. Brown. "David Hume." The Stanford Encyclopedia of Philosophy. Edited by Edward N. Zalta (Summer 2014 edition). URL: <http://plato.stanford.edu/archives/sum2014/entries/hume>.

Napier, Jaime L. and John T. Jost. "Why Are Conservatives Happier than Liberals?" *Psychological Science*, vol.19, no.6 (June 2008): 565-572.

Nedelsky, Jennifer. *Law's Relations: A Relational Theory of Self, Autonomy, and Law*. Oxford: Oxford University Press, 2012.

O'Docherty, Kieran C., Michael K. Mackenzie, Dan Badulescu, and Michael M. Burgess. "Explosives, Genomics, and the Environment: Conducting Public Deliberation on Topics of Complex Science and Social Controversy." *Sage Open* (January-March 2013): 1-17.

Olson, Bradley, J., Satyanarayana Parayitam, and Yongjian Bao. "Strategic Decision Making: The Effects of Cognitive Diversity and Trust on Decision Outcomes." *Journal of Management*, vol. 33, no.2 (April 2007): 196-222.

Ostrom, Elinor. "A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997. *American Political Science Review*, vol.92, no.1 (March 1998): 1-22.

— — —. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press, 1990.

Page, Benjamin and Robert Y. Shapiro. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press, 1992.

Page, Scott E. and Lu Hong. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences*, vol.101, no.46 (November 2004): 16385-16389.

Page, Scott E. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton: Princeton University Press, 2007.

Parkinson, John. *Deliberating the Real World: Problems of Legitimacy in Deliberative Democracy*. Oxford: Oxford University Press, 2006.

Paxton, Joseph M. and Joshua D. Greene. "Moral Reasoning: Hints and Allegations." *Topics in Cognitive Science* vol.2, (2010): 511-527.

Paxton, Pamela and Sheri Kunovich. "Women's Political Representation: The Importance of Ideology." *Social Forces*, 82.1 (2008): 87-113.

Petty, Richard E. and John T. Cacioppo. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer Series in Social Psychology, 1986.

———. *Attitudes and Persuasion: Classic and Contemporary Approaches*. Dubuque, Iowa: Brown, 1981.

Petty, Richard E. and Stephen M. Smith. "Elaboration as a Determinant of Attitude Strength: Creating Attitudes that are Persistent, Resistant, and Predictive of Behavior." In *Attitude Strength: Antecedents and Consequences*, edited by Richard E. Petty and Jon A. Krosnick, Ohio: Psychology Press, 1995.

Petty, Richard E. and Duane T. Wegener. "The Elaboration Likelihood Model: Current Status and Controversies." In *Dual Process Theories in Social Psychology*, edited by Shelley Chaiken and Yaacov Trope. New York: Guilford Press, 1999, 41-92.

Pizarro, David A., Brian Detweiler-Bedell, and Paul Bloom. "The Creativity of Everyday Moral Reasoning." *Creativity and Reason in Cognitive Development*. (2006): 81-98.

Plato. *Republic*. Translated by C.D.C. Reeve. Indianapolis: Hackett, 2004 (circa 380 BCE).

Quick, Kathryn and Jodi Sandfort. "Learning to Facilitate Deliberation: Practicing the Art of Hosting." *Critical Policy Studies* (2014): 1-32.

Rawls, John. *A Theory of Justice*. Cambridge: Belknap Press, 1971.

Raz, Joseph. *The Morality of Freedom*. Oxford: Oxford University Press, 1986.

Reykowski, Janusz. "Deliberative Democracy and 'Human Nature': An Empirical Approach." *Political Psychology*, vol. 27 (2006): 323-346.

Riker, William H. "The Two-Party System and Duverger's Law: An Essay on the History of Political Science." *The American Political Science Review*, vol. 76, no. 4 (December 1982): 753-756.

Rostbøll, Christian F. *Deliberative Freedom: Deliberative Democracy as Critical Theory*. New York: SUNY Press, 2008.

Rubin, Mark and Miles Hewstone. "Social Identity, System Justification, and Social Dominance: Commentary on Reicher, Jost et al., and Sidanius et al." *Political Psychology*, vol.25, no.6 (2004): 823-844.

Russell, Bertrand. "The Nature of Truth." In *The Collected Works of Bertrand Russell, Volume 4*, edited by Alasdair Urquhart. London and New York: Routledge, 1994.

Saltzstein, Herbert D. and Tziporah Kasachkoff. "Haidt's Moral Intuitionist Theory: A Psychological and Philosophical Critique." *Review of General Psychology*, vol. 8, no. 4 (2004): 272-282.

Schnall, Simone, Jonathan Haidt, Gerald L. Clore, and Alexander H. Jordan. "Disgust As Embodied Moral Judgment." *Personality and Social Psychology Bulletin*, vol.38, no.8 (August 2008): 1096-1109.

Schneiderhan, Erik and Shamus Khan. "Reasons and Inclusion: The Foundation of Deliberation." *Sociological Theory*, vol.26, issue 1 (March 2008): 1-24.

Scheufele, Dietram and David Tewksbury. "Framing, Agenda Setting, and Priming: The Evolution of Three Media Effects Models." *Journal of Communication*, vol.57 (2007): 9-20.

Schattschneider, Elmer E. *The semisovereign People: A Realist's View of Democracy in America*. Belmont: Wadsworth, 1960.

Schwartz, Barry and Kenneth Sharpe. *Practical Wisdom: The Right Way to do the Right Thing*. New York: Riverhead Books, 2010.

Scruton, Roger. *A Short History of Modern Philosophy*. London: Routledge, 2002.

Searle, John R. "What is an Institution?" *Journal of Institutional Economics*, vol.1, issue 1 (2005): 1-22.

Sherman, David K. and Geoffrey L. Cohen. "The Psychology of Self-Defense: Self-Affirmation Theory," in *Advances in Experimental Social Psychology*, vol.38. Academic Press, 2006: 183-242.

Sniderman, Paul M., Michael G. Hagen, Philip E. Tetlock, and Henry E. Brady. "Reasoning Chains: Causal Models of Policy Reasoning in Mass Publics." *British Journal of Political Science*, vol.16, issue 4 (1986): 405-430.

Spinoza, Benedictus de. *The Ethics*. Translated and edited by G.H.R Parkinson. Oxford: Oxford University Press, 2000 (1677).

Steenbergen, Marco R., Andre Baechtiger, Markus Spornli, and Jurg Steiner. "Measuring Political Deliberation: A Discourse Quality Index." *Comparative European Politics*. Vol.1 (2003): 21-48.

Stronmer-Galley, Jennifer. "Measuring Deliberation's Content: A Coding Scheme." *Journal of Public Deliberation*, vol.2, issue 1 (2007): article 12.

Taber, Charles S., Damon Cann, and Simona Kucsova. "The Motivated Processing of Political Arguments." *Political Behavior*. Vol.31 (2009): 137-155.

Taber, Charles S., Milton Lodge, and Jill Glathar. "The Motivated Construction of Political Judgments." In *Citizens and Politics: Perspectives from Political Psychology*, edited by James H. Kuklinski. Cambridge: Cambridge University Press, 2001: 198-226.

Talisse, Robert B. "Deliberation." In *The Oxford Handbook of Political Philosophy*, edited by David Estlund. Oxford: Oxford University Press, 2012.

Thompson, Dennis F. "Deliberative Democratic Theory and Empirical Political Science." *Annual Review of Political Science*, vol. 11 (2008): 497-520.

Torrance, Ellis P. "Current Research on the Nature of Creative Talent." *Journal of Counseling Psychology*, no.6 (1959): 309-316.

Touré-Tillery, Rima and Ayelet Fishbach. "How to Measure Motivation: A Guide for the Experimental Social Psychologist." *Social and Personality Psychology Compass*, vol.8 (2014): 328-341.

Turner, Jonathan H. *The Institutional Order: Economy, Kinship, Religion, Polity, Law, and Education in Evolutionary and Comparative Perspective*. New York: Longman, 1997.

Tversky, Amos and Daniel Kahneman. "Judgment Under Uncertainty: Heuristics and Biases." *Science*, vol.185, no.4157 (September 27, 1974): 1124-1131.

Uhlmann, Eric L. and Geoffrey L. Cohen. "Constructed Criteria: Redefining Merit to Justify Discrimination." *Psychological Science*, vol.16 (2005): 474-480.

Uleman, James, Leonard S. Newman, and Gordon B. Moskowitz. "People as Flexible Interpreters: Evidence and Issues from Spontaneous Trait Inference." In *Advances in Experimental Social Psychology*, edited by Mark P. Zanna. San Diego: Academic Press: 179-211.

The United States Environmental Protection Agency. "Climate Impacts on Human Health." Epa.Gov. <http://www.epa.gov/climatechange/impacts-adaptation/health.html>. Accessed April 22, 2015.

Van der Straeten, Karine, Jean-Francois Laslier, Nicolas Sauger, and Andre Blais. "Strategic, Sincere, and Heuristic Voting Under Four Election Rules." *Social Choice and Welfare*, vol.35: 435-472.

Wakslak, Cheryl J., John T. Jost, Tom R. Tyler, and Emmeline S. Chen. "Moral Outrage Mediates the Dampening Effect of System Justification on Support for Redistributive Social Policies." *Psychological Science*, vol.18, no.3 (March 2007): 267-274.

Warren, Mark E. "Deliberative Democracy." In *Democratic Theory Today*, edited by Geoffrey Stokes and April Carter. Cambridge: Polity Press, 2002, 173-202.

———. "Citizen Representatives." In *Designing Deliberative Democracy: The British Columbia Citizens' Assembly*, edited by Mark Warren and Hilary Pears. Cambridge: Cambridge University Press, 2008: 50-69.

Wilson, Timothy D. and Jonathan W. Schooler. "Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions." *Journal of Personality and Social Psychology*, vol.60, issue 2 (February 1991): 181-192.

Wyss, Dominik, Simon Beste, and André Bächtiger. "A Decline in the Quality of Debate? The Evolution of Cognitive Complexity in Swiss Parliamentary Debates on Immigration (1968-2014)." *Swiss Political Science Review*, vol.21, issue 4 (December 2015): 636-653.

Young, Iris M. "Difference as a Resources for Democratic Communication." In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg. Cambridge: MIT Press, 1997.

Young, Iris M. *Justice and the Politics of Difference*. Princeton: Princeton University Press, 1990.

Zaller, John. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press, 1992.