

*The London School of Economics and Political Science*

**Fellow Strangers:  
The Mirage of Collective Political Agency**  
Kaveh Pourvand

A thesis submitted to the Department of Government of the London School of Economics  
for the degree of Doctor of Philosophy, London, September 2019

## **Declaration**

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

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

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party. I declare that my thesis consists of 91,585 words.

## **Abstract**

This thesis criticises the view that citizens do or potentially could control and shape the political institutions that govern them. Such a view conceives of political institutions as the product of the collective agency of the liberal citizenry. This thesis treats John Rawls' political philosophy as paradigmatic of this view, one very prevalent in contemporary political philosophy. Drawing on the social theory of Friedrich Hayek, it argues that we know too little about one another in a large, impersonal society to meaningfully act together in the pursuit of shared objectives. This extends to our political institutions, which are the product of human action but not design. We liberal moderns must reconcile ourselves to a political world that exceeds our capacity to control and direct it.

## Acknowledgements

A PhD thesis can appear as if its creation was a solitary effort. It never is. One needs help to arrive at the starting line of the journey and subsequently one needs help to reach the finish. The author of this thesis is no exception. I would like to begin by thanking my Mother without whose support in so many different ways over the years I would not be writing these words. She, along with her partner Gary, have been a refuge when times have been tough.

I am grateful for the support and friendship of very many LSE fellows and PhD students over my four years at the LSE – Selina, Jan, Elena, Paola, Takuya, Fergus, Kiwi, Elisa, Diane, Max, Jared, Nimrod, Anahi, Cain, Temi, Antoine, Bernardo, Kate, Lucia, Jakob, David, Diana, Alex, Mollie. I was very lucky to study political theory at the LSE. The Political Theory Group (POTY) is welcoming, philosophically rigorous and intellectually diverse. Early versions of two chapters benefited greatly from the forensic dissection of faculty and PhD colleagues at POTY seminars. I am grateful for their feedback on those occasions and for four years of membership in a welcoming academic community.

It took me a while to settle on the main argument of this thesis. I experimented with a fair few ideas first which didn't pan out. My supervisor Chandran Kukathas was patient enough to indulge my explorations and my regular discussions with him were critical to the development and formulation of the final product. I hope his patience has borne some fruit. I also owe special thanks to my advisor Lea Ypi, Kai Spiekermann and Katrin Flikschuh, each of whom have been generous with both their time and help when I have requested it.

A few years ago, before starting the PhD, I used to meet regularly with my friend James Goddard. We would discuss political philosophy, particularly the value of collective action in political life. James and I disagreed on the issue. I now realise that, in many ways, he remained my interlocutor as I wrote this thesis. And I am grateful. I am grateful also to Nick Cowen, who I met also a few years ago when we worked at a think tank. Nick was about to start his own thesis then. He encouraged me to pursue academia at a time when I was uncertain as to what my future held. He's also been a great friend and support since.

## Contents

Figures.....	7
Thesis Introduction.....	8
Thesis Outline .....	11
Notes on Methodology .....	14
Part I: The Theoretical Framework .....	17
Chapter One: Two Visions of Political Society.....	18
Introduction.....	18
1. Associationalism in Rawls’ Political Philosophy.....	19
2. Associationalism Amongst Other Political Philosophers .....	23
3. Disagreement and Impersonal Societies .....	28
4. Alienation Under the Emergent Account .....	35
5. The Emergent Account and the Function of Justice.....	41
6. The Emergent Account and Unintended Outcomes .....	48
7. Hayek’s Worry .....	50
Conclusion .....	52
Chapter Two: “Rawlsekianism” – a Critique .....	55
Introduction.....	55
1. The Limitations of Hayek’s Critique of Social Justice.....	57
2. The Implicit Monocentricity of the Basic Structure .....	61
3. Against Legal Centralism .....	64
4. The State as a Polycentric Network.....	68
5. Gerald Gaus’ Liberalism .....	77
6. Assessing Gaus’ Argument.....	83
Conclusion .....	86
Part II: The Prospects for Collective Political Agency.....	88
Chapter Three: Descriptive Associationalism.....	89
Introduction.....	89
1. The Joint-Intentional Action Model.....	91
2. The Infeasibility of Joint-Intentional Action Under Large-Scale Settings.....	94
4. Liberal Justice as the Object of Joint-Intentional Action .....	102
5. The Principal-Agent Model .....	105
6. Tacit Knowledge and the Limits of Preference Aggregation.....	107
7. The Inevitability of Political Discretion.....	110
8. The Difficulty of Ex-Post Judgements .....	112
9. On the Value of Democracy .....	115

Conclusion .....	117
<b>Chapter Four: Aspirational Associationalism – The Well-Ordered Society</b> .....	<b>119</b>
Introduction.....	119
1. The Rawlsian Assurance Problem as a <i>Legislative</i> Assurance Problem .....	120
2. Deliberation as a Signalling Mechanism.....	126
3. Voting as a Signalling Mechanism .....	129
4. Outward behaviour Versus Internal Motives.....	134
5. The Necessity of Population-Level Judgements.....	137
Conclusion .....	139
<b>Chapter Five: Aspirational Associationalism – Deliberative Democracy</b> .....	<b>140</b>
Introduction.....	140
1. Small-Scale Limitations of Deliberative Democracy .....	142
2. Scaling Up Deliberative Democracy – The Representation Approach .....	144
3. Scaling Up Deliberative Democracy – Deliberation Day .....	148
4. Scaling Up Deliberative Democracy – Systemic Approaches.....	149
5. Tocqueville’s Worry: the Deliberative Virtues and the Perils of Centralisation ...	154
Conclusion .....	161
<b>Part III: Alienation and the Liberal Self</b> .....	<b>162</b>
<b>Chapter Six: Social Unity and the Well-Ordered Society</b> .....	<b>163</b>
Introduction.....	163
1. Goodness as Rationality.....	166
2. Two Models of a Social Union.....	169
3. The Critique of the Emulation Interpretation .....	175
4. The Critique of the Completion Interpretation.....	179
Conclusion .....	186
<b>Chapter Seven: Alleviating Hayek’s Worry</b> .....	<b>189</b>
Introduction.....	189
1. Outlining Hayek’s Worry .....	191
2. The Endogeneity of Political Expectations .....	193
3. Two Accounts of Liberal Stability .....	198
4. The Feasibility of Emergent Liberal Stability.....	201
5. The Rawlsian Objection to Emergent Stability.....	207
Conclusion .....	216
<b>Thesis Conclusion: The Freedom “Inherent” in Agency</b> .....	<b>217</b>
<b>Bibliography</b> .....	<b>221</b>

## **Figures**

<b>Figure 4.1 – Rawlsian Mutual Assurance Game</b>	<b>122</b>
<b>Figure 4.2 – The Difference Principle Contribution Curve</b>	<b>128</b>
<b>Figure 7.1 – The Prisoner’s Dilemma</b>	<b>212</b>
<b>Figure 7.2 – Mutual Assurance Game</b>	<b>213</b>

## Thesis Introduction

We live in impersonal social orders. We inhabit political associations that have been called “imagined communities” because each member is acquainted only with a miniscule subset of her fellow citizens.<sup>1</sup> We are also members of very large cooperative schemes, principally but not only economic, that extend well beyond national borders. This is historically unprecedented. For most of our history we have inhabited largely autarkic and face-to-face social orders in which we are personally acquainted with the fellow members of our cooperative schemes. Much can be said in favour of this shift towards impersonal social arrangements. Market exchange and the extended division of labour have greatly increased our wealth and also our well-being across a wide range of indicators. The expanding of our social world beyond our immediate tribe means we have many more potential life-options available to us than did our ancestors. Life in impersonal orders have also expanded the scope of our moral concern beyond our immediate kin to a much wider range of human and sometimes even non-human beings.

Moreover, the wealth and wider opportunities that impersonal orders generate are arguably preconditions for the emergence and flourishing of liberal individuality. One might posit that a recognisably liberal order must be an impersonal social order.<sup>2</sup> At the very least, the liberal societies with which we are familiar are impersonal ones. The concern of this thesis is the fate of such impersonal and liberal social orders. For all their benefits, these orders are not free of discontent. They expand immensely the scope of social cooperation. Yet their operation, and particularly the differential outcomes they engender, do not appear grounded in any rational or shared foundation. They are often experienced as the products of an alien power whose actions are arbitrary if not unjust. This lack of collective control also alienates us from our peers in impersonal orders, peers with whom we share a social world but no unifying sense of purpose. In this regard, impersonal social orders stand in tension with a widely held democratic intuition: that the good political society is one where citizens work together in order to fashion their political institutions according to their shared and rational will. This intuition is that the good society is marked by “collective political agency”. The goods that one derives from collective political agency are participation in the exercise of collective control over one’s institutions and social unity with one’s peers via their

---

<sup>1</sup> I refer to benedict Anderson’s (2006) famous work of the same title.

<sup>2</sup> Though one should be much less confident in making the reverse claim that impersonal orders are necessarily liberal.



participation in this shared endeavour. Impersonal orders, then, appear to deprive individuals of these great goods.

In response to this predicament, there seem to be two opposing directions in which political philosophy might travel. The first is to theorise the conditions under which collective political agency may be realised in the modern world. This is the strategy adopted by John Rawls. He once wrote that “the comparative study of . . . well-ordered societies is, I believe, the central theoretical endeavour of moral theory”.<sup>3</sup> A society is well-ordered in Rawls’ sense when “everyone accepts and knows that the others accept the same principles of justice, and... the basic social institutions generally satisfy and are generally known to satisfy these principles”.<sup>4</sup> In other words, a well-ordered society is one where citizens share the same intention as to which kind of political institutions are desirable and work together to achieve it. Rawls always emphasised that just institutions should be attained and maintained through the active agency of the liberal citizenry themselves. It should not be dictatorially or paternalistically imposed upon them. Citizens of the well-ordered society “act willingly so as to give one another justice over time”.<sup>5</sup>

This thesis treats Rawls as the paradigmatic representative of what it terms the associational view of political society. This view assumes that collective political agency is possible in a liberal impersonal order. The association in which it is exercised is the state, through which free and equal citizens come together to collectively manage their own affairs. They form communities that, as David Miller puts it, “do things together, take decisions, achieve results, and so forth”.<sup>6</sup> This understanding of political society is very common if not ubiquitous within contemporary political philosophy.<sup>7</sup> Those who adopt this view are deeply struck by the ethical disagreements of liberal citizens. They worry that ethical diversity might prevent liberal citizens from converging on shared aims or principles on the basis of which they could govern their association. Rawls wrestled with this dilemma his whole career. His hope, at least initially, was that liberal citizens could all affirm his famous two principles of justice whatever other disagreements they might have. At the same time, the recurring criticism of the Rawlsian project has been that it is overly optimistic about the prospect for intellectual

---

<sup>3</sup> (Rawls 1974, 13)

<sup>4</sup> (Rawls 1999, 4)

<sup>5</sup> (Rawls 1996, 142)

<sup>6</sup> (Miller 1997, 24)

<sup>7</sup> E.g. (Miller 1997, 2007); (Stilz 2009); (Pevnick 2014); (Habermas 1996); (Pettit 1997); (Waldron 2004)

agreement under modern conditions.<sup>8</sup> What is noteworthy, however, is that few of these critiques have challenged the underlying model of political society.<sup>9</sup> The alternative offered to the well-ordered society by many of Rawls' critics is still an associational model of political society but one that places greater emphasis on procedures by which citizens can manage their disagreements while they are engaged in the project of collective self-governance.<sup>10</sup>

This thesis takes a different tack. It calls into question the associational model itself. In making this argument, it draws heavily on the liberalism of Friedrich Hayek. The basic claim is that the associational model does not take sufficient account of the deep and pervasive mutual ignorance that exists between liberal citizens under impersonal conditions. Hayek constantly stressed that in an impersonal society, each individual acts on significant amounts of private information unavailable to the rest. I call this Hayek's fragmentation of knowledge thesis. On this view, the outcomes of the interactions of many agents under impersonal conditions are not the result of any shared intention despite being the result of many discrete human actions. That is, they are unpredictable and irreducible to the intentions of the actors causally responsible for them precisely because each acts on their own discrete information not known to the rest. In this sense they are *emergent* outcomes.<sup>11</sup>

They are also *necessarily* emergent outcomes because under impersonal settings mutual ignorance about the mass of one's peers and their circumstances is a hard constraint. Joint-intentional action by a group is possible only when the shared objective and how each person's contribution combines with the contributions of others to achieve that objective is common knowledge between them. Their irredeemable ignorance of one another, however, means that liberal citizens could not harmonise their intentions in this way even if they so wished. We have before us not a compliance problem but a *knowledge problem*. This can be overcome in smaller, proximate groupings but not in large-scale groupings of liberal citizens. If this Hayekian argument is defensible then the associational model is found wanting on epistemic grounds.<sup>12</sup> Political institutions are necessarily the product of human action but

---

<sup>8</sup> For variants of this now well-established realist critique, see Williams (2008), Geuss (2008) and Kukathas (2003).

<sup>9</sup> Kukathas' (2003) account of the liberal archipelago being one notable exception.

<sup>10</sup> Jeremy Waldron (2004) and Thomas Christiano (2008) are examples of theorists who take this position.

<sup>11</sup> An emergent phenomenon "exhibits properties that are different from those displayed by the elements [of which it is composed] taken in isolation." (Lewis 2016, 115)

<sup>12</sup> Similarly, the contemporary literature's obsession with intellectual disagreement is overstated. It is not the sole or even the main barrier to realising collective political agency. Chapter one argues that the claim that the mass of liberal citizens are in a state of deep disagreement with one another is greatly exaggerated.

not any shared intentional design. This is the emergence view of political society to be defended in this thesis.

Another key Hayekian theme of this thesis, one already alluded to in the opening paragraph, is the idea that members of impersonal liberal orders might never be fully reconciled to them. Hayek thought that we have certain moral intuitions that are genetically inherited and difficult to dislodge. These intuitions are adapted to the conditions of the evolutionarily basic small group out of which they evolved. They remain embedded deeply enough in the human psyche to ensure a constant demand in the intellectual market place for ideologies, such as nationalism and socialism, which aspire to govern impersonal orders on the same terms as small groupings. We wish to realise at the scale of an impersonal social order the same sense of collective purpose and control we had in the primordial tribal unit. If Hayek's social theory is correct, however, we cannot have this aspiration met and we have to find a way to make do with this fact. This thesis treats the associational view of politics as a species of the same genus thereof. Its great appeal, particularly as expressed by Rawls, is the promise of reconciling liberal individualism with a shared political project and social unity underpinned by liberal justice. If the arguments of this thesis are successful, the Rawlsian hope of such reconciliation are forlorn. The final chapter and the conclusion will engage with the implications of this.

### **Thesis Outline**

The thesis is divided into three parts and seven chapters. Part one, consisting of chapters 1-2, sets out the theoretical framework. Chapter one is expository. It distinguishes the associational from the emergence model of political society. There are three key differences. Associationalism sees liberal citizens as alienated by intellectual disagreement while the emergence view emphasises mutual ignorance. Associationalism conceives of political institutions as the product of conscious and collective design while the Hayekian view emphasises their emergent nature. Associationalism sees the function of justice as providing mutually acceptable terms by which liberal citizens can manage their political association. The emergent view sees the function of justice as facilitating interaction between strangers who know little of one another and their circumstances. This chapter also outlines Hayek's concern that members of impersonal liberal orders would not be reconciled to them and his

worry that they would be institutionally unstable as a result. Chapter two explains why this thesis opts not to follow some contemporary classical liberal thinkers in trying to reconcile the thought of Hayek and Rawls. The chapter considers two such philosophers, John Tomasi and Gerald Gaus, explaining why this thesis rejects their respective approaches. It presents an interpretation of Hayek's ideas stressing that both basic institutional rules *and* the interactions to which they give rise are emergent phenomena. Emergence runs all the way down.

The second part, consisting of Chapters 3-5, attempts to vindicate the emergent over the associational view of political society. In so doing, this thesis distinguishes between two different forms that the advocacy of associationalism might take. *Descriptive* associationalism claims that *actually-existing* liberal democracies are collective political agents. This view is held by figures such as Anna Stilz, David Miller, Ryan Pevnick and Margaret Moore. Descriptive associationalism is invoked to justify various existing practices of liberal democratic states. Thus, liberal democratic states' actually-existing rights to exclude immigrants or control their territory or to non-interference from foreign powers are said to arise because they can be regarded, in their current form, as purposive political associations composed of free and equal citizens.<sup>13</sup> *Aspirational* associationalism claims that the citizens of a normatively ideal regime *would* constitute a collective political agent. Thus, many deliberative democrats lament the incoherence of public opinion under the prevailing practice of aggregative democracy but believe that if practices of deliberative democracy were to be widely adopted, democratic citizens could *then* form a genuine collective will with which the ship of their political association could be steered. Rawls well-ordered society is an unrealised ideal in which, unlike actually-existing liberal democracies, all citizens share the aim of governing their political association on the basis of his famous conception of justice as fairness.

The important difference between the two is that aspirational associationalism posits idealised citizens with certain motivations that their actually-existing counterparts do not have. The citizens of Rawls' well-ordered society are all motivated to comply with Rawlsian justice. This is not true of all their real-world counterparts. The citizens of an ideal deliberative regime exhibit much greater motivation to engage in political deliberation than do their real-world counterparts. I have posited this distinction to try and ensure that this thesis' critique of associationalism is as fair and thorough as it might be. If one critiques only

---

<sup>13</sup> For examples of these three claims respectively see Pevnick (2014), Moore (2015) and Stilz (2015).

descriptive associationalism, advocates of aspirational accounts might respond that their preferred associational model remains unscathed. If only citizens enjoyed a thicker consensus on justice than they do currently or if only they were prepared to deliberate more often, perhaps *then* collective political agency would be possible. Criticising aspirational associationalism is intended to forestall such an objection. Additionally, it serves this thesis' objective of showing that insufficient motivation or the prevalence of ethical disagreement are not the primary barriers to the realisation of collective political agency. The problem is distinctively epistemic.<sup>14</sup> Chapter three, then, considers and critiques the descriptive associational claim that actually-existing liberal democracies are capable of collective political agency. Chapter four considers the prospects for collective political agency if we assume that all citizens affirm Rawlsian understandings of justice. Chapter five considers its prospects in a society in which practices of deliberative democracy are widely adhered to. In each instance, the core claim of this thesis is defended, namely, that collective political agency is not possible given the mutual ignorance of impersonally related liberal citizens.

The third part focuses on the problem of social unity and reconciliation. One of the reasons why the associational view is normatively attractive is that it promises to give individual lives greater significance than they might do otherwise. Rather than merely pursue my own atomised conception of the good, associationalism makes it possible for me to contextualise my own achievements as part of a wider whole. This argument dates back to Hegel and a particularly compelling contemporary version of it is offered by Rawls in his account of the well-ordered society as a social union. That is subjected to critique in chapter six. The upshot of the argument of chapters 2-6 is to vindicate Hayek's claim that an impersonal liberal order cannot meet our need for collective control and social unity. Chapter seven then presents arguments to alleviate Hayek's worry that liberal institutions would be unstable as a result. Hayek feared that liberal citizens would be alienated from the liberal order because of its failure to provide the goods of collective control and social unity. They would constantly be tempted by collectivist political projects that promise to rectify this. While such projects would realise neither unity nor control, their pursuit would have the unintended consequence of undermining a free society. Chapter seven makes two claims to alleviate this concern. First, the preference for social unity and control may not be as invariant as it assumes. Second, and in any case, a liberal order might prove institutionally robust even if many of its citizens were attracted to collectivist political projects. The thesis concludes with

---

<sup>14</sup> We shall have more to say on this point below in the Notes on Methodology.

some reflections inspired by Michael Oakeshott as to how we might make do in a world where we cannot see our own lives as connecting up to a grand social endeavour.

### Notes on Methodology

Before commencing with the argument, I wish to make some notes on methodology.

In many but not all parts of this thesis we will employ a significant idealisation when considering collective political agency. With some exceptions, compliance problems will be idealised away in the arguments to follow. Such idealisation, however, is employed here for a different purpose than the one found in Rawlsian political philosophy.<sup>15</sup> The aim of Rawlsian ideal theory is to outline a “realistic utopia,” showing what might be possible “under reasonably favourable but still possible historical conditions—conditions allowed by the laws and tendencies of the social world”.<sup>16</sup> The function of idealisation for Rawls is to show what might be feasibly achieved in a future society. Rawls does not think full philosophical agreement is a feasible aspiration. He does not idealise his theory by assuming that such agreement would be attained in the well-ordered society. He does think it feasible, however, to assume that all, or at least nearly all, will affirm and comply with a conception of justice whatever other disagreements they might have. Thus, he adds full compliance as an ambitious but nonetheless attainable ideal.<sup>17</sup> Of course, the idealisation of full compliance has been subject to withering criticism.<sup>18</sup>

However, another function that idealisation might serve is to *explicate distinct necessary conditions*. In other words, idealisation might serve as a useful analytical device instead of

---

<sup>15</sup> I here follow Onora O’Neill (1996, 40–41) in distinguishing between an “abstraction” which *brackets but does not deny* predicates true of the matter under discussion and an “idealisation” which *enhances* the predicates under discussion. Assuming full compliance is clearly an “enhancement” of the characteristics of actually-existing liberal citizens rather than an omission of characteristics they do in fact have.

<sup>16</sup> (Rawls 2001, 4)

<sup>17</sup> “We ask in effect what a perfectly just, or nearly just, constitutional regime might be like, and whether it may come about and be made stable under the circumstances of justice (Theory, §22), and so under realistic, though reasonably favorable, conditions”. (Rawls 2001, 13); c.f. (Rawls 1999, 8)

<sup>18</sup> The problem isn’t simply that full or even close to full compliance seems wildly optimistic under any conceivable institutional regime. The problem is that if, as Hume noted, norms of justice are a solution to our limited beneficence, then the same agents who are motivated enough to comply fully with justice would (presumably) also have enough beneficence not to need justice in the first instance. Jacob Levy (2016b, 316–19) articulates this point well. In any case, this thesis will argue that in a large-scale society justice is primarily a solution to difficulties arising from mutual ignorance rather than ethical disagreement.

showing us a desirable and possible but not-yet-attained state of affairs. For instance, it has often been assumed that a lack of requisite motivation is the main impediment to the realisation of a socialist economy. Socialism would be possible if only we could find the right kind of socialist citizen.<sup>19</sup> The genesis of Hayek's political philosophy was the argument that he and his mentor Ludwig von Mises made in the socialist calculation debate. Hayek and Mises wanted to show that socialism would be impossible even if we assume "strict compliance" on the part of all involved because there would remain unmet a distinct condition also necessary for successful central planning. That condition is knowledge of the opportunity cost of the resources under the planner's control. Thus, both took for granted "strict compliance" in their critique of central planning. Their analysis ruled out the possibility of such planning failing because of corruption or shirking on the part of either the planners or the lower-level managers and workers charged with carrying out the plan. Their motivation in so doing was to show how knowledge problems would hinder central planning even under circumstances when compliance with the plan was not at issue.<sup>20</sup> In idealising in this way, they did not need to assume that it would actually be feasible for a socialist economy to be characterised by full or nearly full compliance. Indeed, as stalwart critics of socialism, neither did entertain such an assumption.

This thesis idealises away compliance problems for the same reason. Just as Hayek's knowledge renders infeasible central planning even with full compliance, so does it render infeasible collective political agency. Thus, chapters four and five take for granted that in the Rawlsian society all are committed to liberal justice and that in a deliberative society all are committed to deliberative practices in order to show the distinctively epistemic barrier to collective political agency would remain even if we grant these assumptions. This methodological stance is adopted purely as an analytical device. It should not be taken to imply that strict compliance is a feasible aspiration or that our thinking about normatively desirable institutions should be premised on such an assumption.<sup>21</sup> This said, we shall relax this assumption where it seems appropriate for the argument. For example, it will be relaxed in chapter three because that chapter considers the claim that actually-existing liberal democratic citizens with their actually-existing motivations are capable of collective political agency.

---

<sup>19</sup> Gerald Cohen's *Why Not Socialism?* (2009) is for the most part predicated on this assumption. Though Cohen does pay some recognition to epistemic problems at the end of the essay.

<sup>20</sup> (Mises 1920); (Hayek 1945)

<sup>21</sup> For the same reason, the thesis also idealises away the problems posed by freedom of movement and secessionist movements for the associational model.

One might also wonder about the relation between the associational model, as it is defined here, and nationalism. A common definition of a “nation” is a grouping of persons who recognise one another as distinct from other groupings and as united by certain cultural characteristics such as language or culture.<sup>22</sup> Nations and political associations differ of course. It is by now a truism in these debates to note that many states contain more than one nation and many putative nations lack a state. Some think that national ties are necessary for political association. On this view, the thick ties of national culture are needed to motivate citizens to participate in political association, particularly when such an association calls upon them to make personal sacrifices.<sup>23</sup> Others think thinner ties involving subscription to shared political principles are sufficient.<sup>24</sup> This thesis does not enter into the fray of this dispute, which centres on aforementioned compliance problems that are mostly idealised away in the argument to follow. It thus says little about nationalism *per se* but offers a critique of collective political agency that would hold regardless of whether the parties in question are co-nationals or not. Relatedly, this thesis seeks to critique the associational model *in general* rather than become embroiled in the debate between statist or cosmopolitans as to whether political association should be bounded to a Westphalian-Weberian state or should take an international form, including possibly a world state.<sup>25</sup> It offers a critique of the associational model that would hold regardless of whether political association is limited to contemporary states or extended beyond their current boundaries. Finally, this thesis is Hayekian in inspiration but it makes no claim to be a philosophical exegesis of Hayek’s own views.<sup>26</sup> It is intended to be a contribution to normative political philosophy that draws upon Hayek’s social theory. Indeed, chapter two will explicitly modify Hayek’s ideas in ways of which he may not have approved.

---

<sup>22</sup> Here I follow the definition of David Miller (1997, chap. 2). I accept this definition for the sake of argument. Whether the actual groupings in the world to which we apply the epithet of “nations” actually meet these criteria is another matter.

<sup>23</sup> E.g. (Miller 1997)

<sup>24</sup> E.g. (Moore 2015)

<sup>25</sup> For an influential defence of statism see Nagel (2005) and conversely for a defence of cosmopolitanism see Held (2007).

<sup>26</sup> On that see Kukathas (1989) and Caldwell (2008)



## **Part I: The Theoretical Framework**

# Chapter One: Two Visions of Political Society

## Abstract

This chapter compares the associational understanding of political society as exemplified by John Rawls and others with the emergent understanding derived from Friedrich Hayek's social theory. The associational view conceives of free and equal citizens fashioning political institutions according to deeply shared objectives. The principal bar to such activity is their ethical disagreement. The remedy is impartial principles of justice to which all citizens can assent whatever else they disagree upon. The happy upshot of this approach is its reconciliation of an impersonal liberal order with social unity and collective control. By contrast, Hayekian liberalism is more impressed by the mutual ignorance of impersonally related liberal agents. Liberal justice facilitates coordination between strangers whose ends are opaque to one another. The outcomes of such coordination do not amount to the realisation of any shared objectives. They are the result of many disparate human actions but not any shared design. This approach denies that a liberal social order can be reconciled with social unity and control. The chapter ends by noting Hayek's worry that liberal institutions would consequently alienate their subjects and prove unstable. The remainder of the thesis will try to vindicate the emergent approach.

## Introduction

This chapter sets out the theoretical framework of the thesis. It is divided into seven sections. The first section details how the associational understanding of politics is reflected in John Rawls' political philosophy. The second details how it is prevalent within the work of other political philosophers such as Jurgen Habermas, Joshua Cohen, Anna Stilz, Margaret Moore, David Miller and Ryan Pevnick. The crux of the associational approach is that collective political agency is possible save for widespread intellectual disagreement, a problem readily soluble with suitably impartial principles of justice. Section three criticises the associational assumption that liberal citizens are in a state of persistent and deep disagreement. It argues negatively in defence of what is argued for positively in section four, where we turn directly to Hayek. That section outlines what I call Hayek's fragmentation of knowledge thesis to suggest that liberal agents are alienated from one another primarily by ignorance rather than intellectual disagreement.<sup>27</sup> As a point of entry into Hayek's thought, I also highlight some

---

<sup>27</sup> I should acknowledge my debt here to Adam Tebble's (2016) insightful paper, "On the circumstances of justice". This chapter and indeed this thesis build on its two core claims: (i) that mutual ignorance rather than disagreement is the primary cause of the "circumstances of justice," and (ii) accordingly that the primary function of justice is to facilitate coordination between mutually ignorant agents. Ignorance rather than disagreement is primary because even if we posit that all agents share the same conception of the good, their ignorance of one another's circumstances precludes their coordination, *absent* rules of justice, to achieve it. The argument of this chapter builds on Tebble's insights,

perhaps unlikely similarities between his concerns and those of Iris Marion Young. Section five details the contrasting role played by justice in the emergence account, namely facilitating coordination between strangers who know little of one another. Section six emphasises that on the emergence account we cannot ascribe collective intentionality to large-scale outcomes because they are properly the result of human action but not design. Section seven closes the chapter by highlighting Hayek's worry that a liberal order would be unstable precisely because it cannot meet our deep desire for collective control and social unity.

### 1. Associationalism in Rawls' Political Philosophy

The introduction identified three components to the associational model of political society:

- i. Liberal citizens are alienated by intellectual disagreement
- ii. The function of justice is to secure agreement
- iii. Political institutions are the product of the shared agency of liberal citizens

All three features are present in John Rawls' political philosophy. We shall treat them in turn. That the first two features are present in his thought is uncontroversial enough that a brief treatment will be sufficient. Rawls conceives of political society as a "fair system of cooperation over time from one generation to the next".<sup>28</sup> Rawls says it should be governed by a public conception of justice for the well-known reason that liberal citizens are divided over the good. According to *A Theory of Justice*, one of the subjective circumstances giving rise to the need for justice is the fact that "there exists a diversity of philosophical and religious belief, and of political and social doctrines".<sup>29</sup> In the later formulation of *Political Liberalism*, Rawls similarly says that we turn to a "political" conception of justice because a "diversity of reasonable comprehensive religious, philosophical and moral doctrines. . . is a permanent feature of the public culture of democracy".<sup>30</sup> The remainder of this section will focus on showing how the third component is a feature of Rawls' thought.

---

particularly so in two ways. First, if we take mutual ignorance seriously, then the characterisation of liberal citizens as in a state of disagreement is exaggerated if not mistaken. For they will not know enough about one another's views to be so described. Second, it provides a typology of the different ways in which justice facilitates coordination between epistemic strangers.

<sup>28</sup> (Rawls 2001, 4)

<sup>29</sup> (Rawls 1999, 110)

<sup>30</sup> (Rawls 1996, 36)

The subject of Rawlsian justice is the “basic structure” of society. This refers, of course, to “the way which the main political and social institutions of society fit together, and the way they assign basic rights and duties and regulate the division of advantages that arises from social cooperation over time”.<sup>31</sup> Examples of major social institutions are “the legal protection of freedom of thought and liberty of conscience, competitive markets, private property in the means of production, and the monogamous family”.<sup>32</sup> These institutions coalesce into a structure the operation of which has a profound effect on the life chances of different individuals. This is why they are the appropriate subject of justice. We shall have much more to say about this in chapter two. For now we can note that Rawls must implicitly accept the following claim to make the basic structure the primary subject of justice. We have the requisite knowledge and capacity to intervene upon these institutions understood as an interconnected *system* or *structure* so as to make the outcomes they engender accord with some desired objective i.e. a conception of justice.<sup>33</sup> In other words, it is in principle possible to design institutions holistically to accord with a normative blueprint. The remaining question, then, is *which* blueprint.

For Rawls, the answer is his two famous principles of justice. In the well-ordered society, everyone accepts and knows all others accept the same principles. He includes this stipulation to secure institutional stability of a certain kind. He writes:

The problem of stability is not that of bringing others who reject a conception to share it, or to act in accordance with it, by workable sanctions, if necessary, as if the task were to find ways to impose that conception once we [i.e. Rawls and his readers] are convinced it is sound.<sup>34</sup>

Instead,

[c]itizens should act willingly so as to give one another justice over time. Stability is secured by sufficient motivation of the appropriate kind acquired under just institutions.<sup>35</sup>

On Paul Weithman’s recent influential reading, Rawls opts for *inherent* as opposed to *imposed* stability.<sup>36</sup> Inherent stability is secured when liberal citizens have internalised the principles of justice such that they are intrinsically motivated to comply with them rather than on the basis of external incentives or sanctions. I pay my taxes because I affirm the

---

<sup>31</sup> (Rawls 1999, 7)

<sup>32</sup> (Rawls 1999, 6)

<sup>33</sup> Chapter two will argue that no agent or agency actually has such a capacity. Therefore, assuming ought implies can, the basic structure cannot be the subject of justice.

<sup>34</sup> (Rawls 1996, 143)

<sup>35</sup> (Rawls 1996, 142)

<sup>36</sup> (Weithman 2011, chap. 2)

difference principle and not because I fear the taxman knocking on my door. Inherent stability means that just institutions are secured and realised through the *freely-willed* actions of liberal citizens. Imposed stability – the securing of compliance through “workable sanctions” – is the approach to political philosophy Rawls associates with Hobbes. Rawls’ respect for the autonomy of liberal citizens leads him to prefer inherent stability. A conception of justice should not be imposed on recalcitrant subjects. A just society can be so described because its members have freely-willed it to be so. One might think it sufficient for Rawls to restrict the first condition to stating solely that everyone should accept the principles of justice. The second clause stipulating their awareness that *everyone else* also accepts them might appear unnecessary. However, one person on their own cannot create or secure a just society. It is necessarily a shared endeavour. And a shared endeavour gets off the ground only when the participants can signal to one another their commitment to it.

Were I the only good Rawlsian willing to pay his taxes and vote for Rawlsian principles in a society full of resolute anarcho-capitalists, I would have no chance of realising Rawlsian justice. The success of my contribution to the creation of a just society is contingent on my fellow peers making their own contributions. Thus, I need assurance that they too are willing to contribute to make worthwhile my own contribution. What is true of me is true of all others. It is a truism of the literature on joint-intentional actions that no cooperative activity can get off the ground unless the participants’ mutual willingness to participate is common knowledge between them.<sup>37</sup> Supposing liberal citizens can solve this mutual assurance problem, a topic to be explored in chapter four, we can say the following of the well-ordered society. The operation of its major institutions accords with a conception of justice that is not simply subject to the approval of its citizens. Such institutions continue to be just because liberal citizens jointly share the intention that they be maintained. It is in this sense that the well-ordered society is the product of the collective political agency of its members.

Rawls’ associationalist leanings are also present in his international theory, as expounded in *The Law of Peoples*. Rawls there posits that whole nations or *peoples* have a rational sense of their own interests. They wish to “protect their political independence and their free culture with its civil liberties, to guarantee their security, territory, and well-being of their citizens”.<sup>38</sup> Their rational conduct is “organised and expressed in their elections and votes, and the laws and policies of their government”.<sup>39</sup> On Stephen Macedo’s interpretation, Rawls

---

<sup>37</sup> We shall look at this literature in more detail in chapter three.

<sup>38</sup> Rawls (2003, 34)

<sup>39</sup> Rawls (2003, 25)

formulates his theory in this way to protect different national preferences. We might understand France and the United States, for instance, as making different collective “choices” with respect to work, leisure and economic growth which need to be respected.<sup>40</sup> By the same token, however, “peoples” should be held responsible for their decisions. Rawls makes a hypothetical comparison between an industrialised country which develops after it “decides” to increase “its” savings rate and a “pastoral and leisurely” country that did not make such a decision. He then suggests that it is intuitively clear that, under the assumption that each people is free and responsible and able to make their own decisions, that the industrialised country should not be taxed to give funds to the pastoral one.<sup>41</sup> Rawls views in the international sphere is the counterpart of the claim in his domestic theory that a society’s basic structure is the product of the freely given and jointly coordinated actions of liberal citizens. It follows intuitively from this premise that these citizens are engaged in a form of collective agency deserving of normative recognition and for whose consequences they ought to bear responsibility.<sup>42</sup>

The introduction also identified social unity and collective control as the two great goods secured by political life on the associational view. These positions are found in Rawls’ political philosophy also. First, insofar as the basic structure of society is maintained by the willed intentions of liberal citizens, they inhabit a political and social world of their own construction. Rawls gives this idea a Kantian-Rousseauvian construal when he says that members of the well-ordered society attain their *full autonomy*. That is, the basic structure of their society accords with general principles that they would give themselves were they to adopt the autonomous (in the Kantian sense) position of the original position. When they obey the law or vote according to liberal justice, they act as self-legislators.<sup>43</sup> Second, citizens achieve social unity insofar as they endorse and create the just basic structure *together* on the basis of principles to which they all assent. In the third part of *A Theory of Justice*, Rawls characterises the well-ordered society as a “social union” in which “the successful carrying out of just institutions is the shared final end of all the members of society, and these institutions are prized as good in themselves”.<sup>44</sup> He goes on to say that “this collective

---

<sup>40</sup> (Macedo 2004, 1278)

<sup>41</sup> (Rawls 2003, 117)

<sup>42</sup> However, as I said in the introduction, associationalists don’t need to be statist necessarily. They can be cosmopolitans. It depends on whether the associationalist thinks the relevant political association should be a nation-state or a world state of some kind.

<sup>43</sup> (Rawls 1996, 77)

<sup>44</sup> (Rawls 1999, 462)

activity... must be experienced as a good” and that the “public realization of justice is a value of community”.<sup>45</sup>

## 2. Associationalism Amongst Other Political Philosophers

This thesis characterises Rawls as an aspirational associationalist. The referent of the well-ordered society and the law of peoples are hypothetical members of idealised regimes and not the actual liberal democratic polities with which we are familiar. Deliberative democratic theorists can also be so characterised.<sup>46</sup> They share the Rawlsian assumption that disagreement is the main barrier to collective political agency and that the identification of impartial principles of justice can help overcome this problem. On their view, however, what is required is not so much a society-wide commitment to a particular conception of justice but a commitment to deliberative procedures with which such impartial principles may be identified. What matters is that citizens are willing and able to enter political deliberation the discursive rules of which are premised on the idea that all participants are free and equal.<sup>47</sup> Deliberative democrats such as Jürgen Habermas and Joshua Cohen have formulated similar listings of what the appropriate discursive conditions should be.<sup>48</sup> The basic idea is that the participants should aim for a rational consensus justifiable to all of them understood as free and equal.

This consensus then constitutes the public will on the basis of which their shared polity should be arranged and governed. Hence, in such an ideal deliberative regime, where the motivation to engage in political deliberation is widespread and a rational consensus on justice is attained, authentic collective political agency would be possible. By contrast, deliberative democrats tend to think that the prevailing practice of aggregative democracy is a poor method of eliciting the general will.<sup>49</sup> Deliberative democrats tend also to believe that, apart from collective control, deliberative participation can help foster the values of community and social unity. For instance, Joshua Cohen claims that the discursive

---

<sup>45</sup> (Rawls 1999, 463); we shall investigate Rawls’ idea of a social union in chapter six.

<sup>46</sup> (Jürgen Habermas 1983, 1996); (Cohen 1997); (Gutmann and Thompson 1998); (Mansbridge et. al. 2012); (Ackerman and Fishkin 2004)

<sup>47</sup> Habermas (1995) famously criticises Rawls’ original position as “monological”. Habermas employs this term to highlight the fact that the original position does not leave scope for any real deliberation about what justice requires. Rawls effectively stipulates principles to which he thinks liberal citizens *would* agree. For a deliberative democrat like Habermas, one cannot second guess the object of consensus in this way. It has to be arrived at through actual deliberation.

<sup>48</sup> (Cohen 1997, 74–75); Habermas (1983)

<sup>49</sup> E.g. (Cohen 1997, 75); Habermas (1996, 506)

requirement that deliberation be justifiable to all “connects” with the value of community.<sup>50</sup> He suggests that deliberative participation shapes citizens’ identity and preferences so as to be aligned with the common good.<sup>51</sup> We shall consider the claims of deliberative democracy in detail in chapter five.

There are also many philosophers who adopt a descriptive associational view. These philosophers believe we can characterise actually-existing liberal democracies as associations of free and equal citizens who have together shaped and determined the institutions under which they live. They do so to justify claims made by existing liberal states. The remainder of this section will survey their ideas. Anna Stilz distinguishes between a “taker” view of political institutions conceiving of citizens primarily as the beneficiary of rights and other entitlements and a “maker” view under which they are the active creators of the political and social institutions under which they live.<sup>52</sup> The exercise of such maker freedom represents something greater than the exercise of mere private rights. Citing Hegel, she writes:

Freedom additionally requires that individuals who sustain state institutions together experience this activity as an expression of themselves, not as something that they are coerced into performing by an alien power. If citizens attain subjective freedom, they will see their state as a creation of their own free cooperation, not as an institution of subjugation.<sup>53</sup>

The political, social and economic institutions to which citizens are subject are not mere “happenings” beyond their control and to which they merely acquiesce. Rather they endure because all citizens intend that they do and each freely plays her part in jointly-coordinated activity to realise this shared intention.<sup>54</sup> Thus for Stilz, as for Rawls, political institutions should be the product of the shared agency of liberal citizens. Unlike Rawls, however, she doesn’t offer her account as a description of the relations enjoyed by the hypothetical citizens of an idealised well-ordered society. Hers is a description of the relations enjoyed by real-world citizens of actually-existing liberal democratic regimes. In providing such an account she hopes to justify certain moral claims that such regimes make. Hence, claims Stilz, if actually-existing liberal democratic regimes are sustained by the shared and active agency of their citizenry, and such shared agency constitutes a morally valuable relationship with

---

<sup>50</sup> (Cohen 1997, 74–75)

<sup>51</sup> (Cohen 1997, 76-77)

<sup>52</sup> (Stilz 2015, 7–8)

<sup>53</sup> (Stilz 2015, 12)

<sup>54</sup> (Stilz 2015, 13-14)



which outsiders should not interfere, then the rights of sovereignty and territory claimed by such regimes are justifiable.<sup>55</sup>

In her defence of territorial rights, Margaret Moore identifies three conditions under which a group of persons count as a collective agent or “people” that can legitimately claim territorial rights over a geographical space. First, they must conceive of themselves as a group “in terms of either being engaged, or desiring to be engaged, *in a common political project* and they are mobilized in actions orientated towards that goal”. Second, they “must have the capacity to establish and maintain political institutions, through which they can exercise self-determination”. Third, they must have a history of political cooperation with one another.<sup>56</sup> Moore is keen to stress that the terms of this common project must be *political* such that “people from a wide range of backgrounds and with different conceptions of the good are nevertheless prepared to cooperate and compromise in the creation of and maintenance of rules to govern their lives”.<sup>57</sup> In other words, the shared association should be governed by political principles of justice to overcome the problem of disagreement. Like Rawls, Moore too emphasises the twin benefits of collective control and social unity enjoyed by those fortunate enough to be members of a political association. She writes:

People who exercise collective self-government have the institutional mechanisms to shape the conditions of their existence, and their future together, and are thereby more autonomous—or experience a different (collective) dimension of autonomy than is involved in most liberal accounts of autonomy, which are mainly focused on protecting autonomy through protecting the individual private sphere.<sup>58</sup>

Further, the civic relationship has a value comparable to family relationships. The family relationship is good for children not simply because it instrumentally provides them with generic goods such as food and clothing and so on. It is valuable also because children enjoy a relationship of love and care with their family. The *manner* in which the generic goods are secured is of intrinsic value.<sup>59</sup> *Mutandis mutatis*, citizens enjoy a comparable relationship with one another, one that has an intrinsic value distinct from the generic goods that political life provides them, such as having their rights respected.<sup>60</sup> As with Stiliz, Moore offers her account to describe actually existing liberal democratic states in order to justify certain moral claims that they and/or their citizens make. Moore too is concerned with territorial claims.

---

<sup>55</sup> (Stiliz 2015, 8-12)

<sup>56</sup> (Moore 2015, 50–51); original emphasis

<sup>57</sup> (Moore 2015, 53)

<sup>58</sup> (Moore 2015, 66)

<sup>59</sup> (Moore 2015, 48)

<sup>60</sup> (Moore 2015, 63)

On her account, liberal democratic states can legitimately lay claim to the territory under their control because they are composed of citizens engaged in the morally valuable relationship of collective self-governance and such a relationship would not be possible without the requisite territorial control.

In a series of works, David Miller has defended the claim that “[nations] are communities that do things together, take decisions, achieve results, and so forth”.<sup>61</sup> We can, he thinks, speak of people having a shared will under reasonably functioning democratic institutions. For then “the political decisions that are taken will embody to a greater or lesser extent the articulated beliefs and attitudes of the nation in question”.<sup>62</sup> Under such circumstances the “policies pursued by the state can reasonably be seen as policies for whose effects the citizen body as a whole is collectively responsible, given that they have authorized the government to act on their behalf in a free election”.<sup>63</sup> Like Moore, he draws an analogy between the relationships we have with our families and friends and those we have with our co-nationals. Just as we have strong and partial ties with families and friends because they reflect “existing [and particularistic] commitments, relationships, and loyalties,” so do we have such ties with our co-nationals.<sup>64</sup> Indeed, the ties of national history can be very weighty for Miller. Co-nationals have duties towards one another and also towards their descendants because they, in turn, have benefited from the sacrifices of their “forefathers” who “toiled and split their blood to build and defend the nation”.<sup>65</sup>

Like Rawls, Miller wants to identify the conditions under which citizens would be motivated to take part in political association. However, Miller has more of a communitarian bent than Rawls and other liberals such as Moore or Stilz. He thinks that intellectual agreement on an abstract conception of justice is at best necessary but not sufficient to motivate citizens to participate in political association. The aforementioned national ties are also necessary.<sup>66</sup> Though Miller also stresses that the shared national culture should be civic and open to the plurality of cultural groups and opinions that one finds in contemporary states.<sup>67</sup> So he still

---

<sup>61</sup> (Miller 1997, 24)

<sup>62</sup> (Miller 2007, 127)

<sup>63</sup> (Miller 2007, 128)

<sup>64</sup> (Miller 1997, 57)

<sup>65</sup> (Miller 1997, 23)

<sup>66</sup> (Miller 1997, 57–58); as stated already in the introduction, this thesis will say little about nationalism per se when considering the prospects for associationalism. If Miller is right, then nationalism can only solve the motivational barriers to collective political agency whereas this thesis is focused on epistemic barriers. For a trenchant critique of Miller’s position, see Abizadeh (2007).

<sup>67</sup> (Miller 1997, chap. 6)

shares the associational instinct that cognitive diversity or disagreements over ways of life are the main obstacles to collective political agency. As with Stiliz and Moore, Miller's account is offered as a description of actually-existing states in order to justify pre-existing moral claims and practices. For instance, according to Miller, it is in virtue of the fact that the people control the state under reasonably functioning democratic institutions that we can justify the oft-found belief that all co-nationals are responsible for the actions of their government.<sup>68</sup>

Ryan Pevnick argues that liberal democratic states are purposive associations that acquire property rights over their territory via their shared labour. He believes his quasi-Lockean account to be already implicit in Rawls' characterisation of the well-ordered society as a cooperative venture between citizens. The citizenry act together to create public goods – “basic infrastructure, defense, the establishment and maintenance of an effective market, a system of education, and the like.”<sup>69</sup> These arise though “the coordinated decisions, labor, and contributions of members [of the association]”.<sup>70</sup> The members are then collectively entitled to associative ownership of these institutions. Pevnick connects his thought with the aforementioned Rawlsian concept of a social union, arguing that our contributions to a framework of publicly provided goods heightens our own personal achievements, for subsequently “each of us accomplishes more and lives more broadly than any of us could on our own”.<sup>71</sup> Again, we see an analogy with the family cropping up. Just as the inheritor of a farm claims legitimate title to it, so do the current generation of liberal citizens claim legitimate title to the public goods bequeathed to them by their predecessors.<sup>72</sup> Like the three previous authors, Pevnick intends his account as a description of the practices of actually-existing liberal democratic states to justify their right to regulate immigration. On Pevnick's model, the citizenry of a nation have acquired a property right over their territory via their coordinated and intentional labour. Just as any property-owner has at least a defeasible right to determine who to admit on her property, so do liberal citizens have a qualified right to determine who to admit into their political association.<sup>73</sup> We shall assess the prospect for descriptive associationalism in chapter three.

---

<sup>68</sup> (Miller 2007, chap. 5)

<sup>69</sup> (Pevnick 2014, 38)

<sup>70</sup> (Pevnick 2014, 53)

<sup>71</sup> (Pevnick 2014, 38)

<sup>72</sup> (Pevnick 2014, 37–38)

<sup>73</sup> (Pevnick 2014, chap. 1)

### 3. Disagreement and Impersonal Societies

Our outline of the associational view of political society as found in the work of Rawls and other political philosophers is complete. Section four will outline the alternative emergent understanding derived from Hayek's work. To motivate that discussion, this section criticises the common associational assumption that liberal citizens are alienated primarily by intellectual disagreement. By "alienation" I mean simply that which prevents liberal citizens from cooperating politically on the basis of shared ends. Rawls' inquiry was always premised on the assumption that intellectual disagreement was the relevant bar. Finding a solution to the problem vexes him his whole career—so much so that it inspires the change in his justificatory strategy between *A Theory of Justice* and *Political Liberalism*. Fearing that the former draws too heavily on Kantian and Millian understandings of autonomy that not all liberal citizens could find acceptable, he reframes his argument for justice as fairness as the articulation of ideas already implicit in the public political culture of liberal democracies and to which, he now hopes, liberal citizens would have reasons to assent despite all their other philosophical differences. In his later work he also allows for the possibility that cognitive disagreement might extend to justice itself.<sup>74</sup>

By the end, he offers not so much a theory of justice as a theory of the appropriate kinds of *public reasons* that citizens may cite in exercising power over one another as they manage their political association. This project is continued by scholars working in the tradition of public reason liberalism.<sup>75</sup> Great intellectual effort has thus been invested by Rawls and those following in his wake to show how the problem of disagreement in political life can be overcome. Many realist scholars remain unconvinced, believing that our ethical diversity simply precludes the deep consensus implicit within the well-ordered society. On their view, we have to theorise about ongoing political life in a manner that accounts for much more disagreement than Rawls' theory can cope with.<sup>76</sup> What nevertheless unites Rawls and his critics is the view that ethical disagreement between liberal citizens is the central datum from which political theorising should begin. This section will cast doubt on this assumption.

Quite simply, I claim that our ignorance of the views of most of our fellow citizens makes it very difficult to know whether we agree with them or not. Therefore when Rawls refers to a "just and stable society of free and equal citizens, *who remain profoundly divided by*

---

<sup>74</sup> Rawls' justice pluralism is evident in, for example, "The Idea of Public Reason Revisited" (1997).

<sup>75</sup> Though they disagree as to how to proceed. For two contrasting accounts see Gaus (2012) and Quong (2011).

<sup>76</sup> Bernard Williams (2008) and Raymond Geuss (2008) offer two prominent realist accounts.

*reasonable religious, philosophical and moral doctrines,*” he is mischaracterising the relationship that obtains between members of large-scale and impersonal social orders.<sup>77</sup> To clarify, I do not deny that such orders are marked by a plurality of ethical views and their citizens have some awareness of this social fact. Nor do I deny that there are ethical cleavages of political significance in contemporary liberal democracies, such as the abortion controversy in the United States. I also accept the thrust of Rawls’ account of “reasonable disagreement,” that different human agents non-culpably reach divergent conclusions when considering complicated and difficult issues. However, it goes too far to infer from these points that the mass of impersonally related liberal citizens are in a relation of “deep doctrinal conflict”<sup>78</sup> that is “irreconcilable”<sup>79</sup> and of “absolute depth”.<sup>80</sup>

To explicate this claim further, let us look more closely into what Rawls means by reasonable disagreement. Rawls believes that liberal citizens will have reasonable disagreements over what he terms “comprehensive doctrines”. If a belief is a view on a single subject matter, a comprehensive doctrine is an integrated and consistent set of beliefs extending to “issues of value in human life, and ideals of personal character, as well as ideals of familial and associational relationships, and much else that is to inform our conduct”.<sup>81</sup> Such a doctrine should cover the “major religious, philosophical, and moral aspects of human life” and provide guidance as to how to deal with value conflicts.<sup>82</sup> The claim of reasonable disagreement holds that individuals who exercise their power of reasoning will arrive at differing comprehensive doctrines that are incommensurable with one another and which render incompatible judgements on various matters. They will arrive at these divergent doctrines non-culpably due to burdens of judgement that encumber all rational agents when deciding upon complex matters.<sup>83</sup>

The first difficulty with this account stems from Rawls’ own concession that most “people’s religious, philosophical and moral doctrines are not seen by them as fully general and comprehensive, and these aspects admit of variations of degree”.<sup>84</sup> Indeed, it seems that only a philosopher of a certain kind will form a comprehensive doctrine in Rawls’ sense. (I

---

<sup>77</sup> (Rawls 1996, 4); emphasis added.

<sup>78</sup> (Rawls 1996, xxviii)

<sup>79</sup> (Rawls 1996, 441)

<sup>80</sup> (Rawls 1996, xxviii)

<sup>81</sup> (Rawls 1996, 13)

<sup>82</sup> (Rawls 1996, 13)

<sup>83</sup> For instance, the evidence may be conflicting or the assessment may involve several considerations with no obvious way of weighting them. The full list of such burdens is provided in Rawls (1996, 56–57).

<sup>84</sup> (Rawls 1996, 160)

freely confess to lacking one myself.) Clearly, if citizens lack an opinion on a topic, they can neither be described as in agreement or disagreement on it. Some proof for this contention comes from the fact that the first thing taught to political philosophy students are the various positions that one can take in the discipline. The difference between consequentialist and deontological ethical positions might be a bread-and-butter example. The students have to be taught these positions precisely because most people don't have a determinate opinion either way *before* studying the discipline. Even if we restrict our attention to political beliefs, empirical evidence suggests that many if not most citizens do not subscribe to a formal political philosophy.<sup>85</sup> The empirical evidence on voter ignorance of basic political facts suggests also that many voters are simply apathetic about politics.<sup>86</sup>

This is not the main difficulty however. Suppose that all liberal citizens did take the time to form comprehensive doctrines and become well-informed about politics. Perhaps such circumstances would obtain in a future Rawlsian society. Granting that under such circumstances liberal citizens would form comprehensive doctrines that *differ in content*, it would still not follow that they would be in a state of *agreement or disagreement*—at least if the latter terms are understood in the relevant intersubjective sense. It would prove helpful here to distinguish between two senses in which the notion of agreement is used. The statement “Rosie and Jim agree that proposition X is true” doesn't necessarily mean that either are *first-personally* aware that the other also believes X or indeed of the other person's existence. It could simply mean that they have both *independently* concluded that X is true. The statement “Aristotle and Confucius agree that excellence of character is important,” for example, can be made plausibly only in this sense since neither could have been aware of the other's existence. However, the statement “Rosie and Jim agree that proposition X is true” could also be interpreted as indicating that Rosie and Jim both believe that X is true *and* that each knows that the other believes that X is true. They could each become aware of this after having independently determined the truth of X or perhaps after having deliberated with one another to arrive at that conclusion.

In other words, the terms “agreement” or “disagreement” could be used to refer solely to a correspondence of views (or lack thereof) or with the extra requirement that the parties being referred to it are *cognisant* of that fact. Our normal usage of these terms presumes both conditions are in place. So does the particular phrasing that two or more people are in

---

<sup>85</sup> The classic study of voters' lack of any discernible ideology remains that of Phillip Converse (2006).

<sup>86</sup> For a review of the empirical evidence on voter ignorance, see Somini (2013, chap. 1).

a *state* of agreement or disagreement. Certainly, political theorists presume both conditions are met when they characterise liberal citizens in such a manner. When Rawls says liberal citizens are profoundly “*divided*” by various doctrines, we can only take him to mean that they are aware that their opinions differ.<sup>87</sup> It is hard to see how a group of persons could be divided by intellectual differences that they don’t know they have. Thus, if liberal citizens are to be so described as disagreeing in the relevant intersubjective sense, they need to have mutual knowledge of one another’s views. If they are to be described as in a state of disagreement that is “*irreconcilable*” and of “*absolute depth*” they need a *high degree* of mutual knowledge of one another’s views.<sup>88</sup> However, in an impersonal society each citizen will have such concrete knowledge of the views of only a subset of their peers – friends, family and co-workers say. Given the limited knowledge that each individual citizen has of most of her peers, one can characterise liberal citizens *en masse* as being in a state of agreement or disagreement only to a very limited extent.

Now, the Rawlsian might concede this but object that it is beside the point with which Rawls is ultimately concerned. What matters is not whether citizens know all the important differences in one another’s comprehensive doctrines but whether they know of those *policy areas* with respect to which their differing comprehensive doctrines lead them to disagree. For instance, the feminist does not need to know all the intricacies of Catholic doctrine to know that he will disagree with many Catholics about the permissibility of abortion. (And vice-versa.) Rawls’ characterisation of liberal citizens as in a state of disagreement *en masse* rests only on their knowing the *politically salient* differences in their comprehensive doctrines. The difficulty with this objection is that even with respect to political preferences, knowledge of what the mass of our fellow citizens think is not easy to come by. One prime facie plausible mechanism by which we might learn the political preferences of our fellow citizens is their voting patterns. Thus, we shall respond to this objection by demonstrating that voting patterns reveal much less about citizens’ political preferences than one might think.

---

<sup>87</sup> (Rawls 1996, xxv); emphasis added.

<sup>88</sup> The thought here being that I need to know a great deal about another citizen’s views to be able to say that our disagreements are as stark as the Rawlsian story intimates i.e. we don’t just disagree about, say, a particular policy but also on the deeper philosophical and even metaphysical premises that inform our policy positions. But absent substantive knowledge of the actual comprehensive doctrines of most of my fellow citizens, how can I know that my disagreements (if indeed we do disagree) with them extend to the “*absolute depth*” of which Rawls speaks?

The fact that 42.4% and 40% of the United Kingdom electorate voted for the Conservative and Labour parties in the 2017 General Election might be taken to imply that 42.4% of the UK electorate favour the contents of the Conservative Party manifesto and 40% the Labour Party manifesto. This inference is dubious enough in the real world where it is safe to say that voters are not always fully knowledgeable about the policies of the parties or candidates for whom they vote. However, even if we posit well-motivated voters who carefully read manifestos, the inference would still be mistaken. For voters do not always agree with every policy of the candidates they vote for. It is quite conceivable that a party voter would agree with only some of its policies and disagree with others but still vote for it because they conclude that the benefits of implementing the policies they favour outweigh the costs of implementing those they don't. Given the highly limited range of choices in the political market place and the heterogeneity and plurality of the citizens governed by liberal states, it seems that many voters would have to compromise in this manner when deciding how to vote.

If so, this greatly limits the utility of voting patterns as a heuristic to assess the views of one's fellow citizens. When our well-informed and motivated voters opt for a party with a particular manifesto, the most we can say for sure is that they approve *some* of the policies it contained. We do not know which, if any, of those policies the voter disagreed with but judged to be a price worth paying for the benefits of the policies they did favour. Nor do we know the relative strengths of their preferences for or against a particular policy. A real-world example might prove useful here. Until recently the Republican Party of the United States had campaigned strongly on a platform of free trade. If we naively took voting patterns as an effective proxy of voters' political views, then we might have concluded that Republican Party voters were also staunch supporters of free trade. Yet, Donald Trump won both the 2016 Republican Party presidential primaries and the presidential election on an explicitly protectionist platform, suggesting that free trade did not actually rank highly in the preferences of many Republican Party voters in the recent past. Indeed, perhaps for many, their optimal policy all along was protectionism. It looks as if they voted for the Republican Party in the past because of its campaigning for other issues that they ranked higher in importance than free trade, a policy with which they might well have disagreed.

It might be objected that this problem is attenuated to the extent that the policies offered by a political party entail one another. Two or more policies entail one another if supporting one is itself reason to support the others. The greater the degree to which the policies of a party entail one another in this way, the more does the decision to vote for them reveal



about the political preferences of their voters. Thus, suppose that a political party runs a policy programme consisting of 26 policies A-Z that entail one another in this sense. Then we might infer, from the mere fact of their voting for them, that the party's voters supports all 26 policies. The mutual entailment assumption means we can rule out the possibility that they trade-off the benefits of favoured against the costs of disfavoured policies. Thus it allows us to infer more from their voting decision than if we allowed for such trade-offs. While without the assumption, the most we can say for sure is that a party voter supports *some* of its policies. With the assumption, we can say that the voter supports *all* the policies.<sup>89</sup>

The trouble is that it is very hard to think of how policies can be mutually entailing in this way. Perhaps they would be so if they *logically* entailed one another. Policy A and B would logically entail one another if I could affirm A and reject B only on pain of logical contradiction.<sup>90</sup> But policy platforms do not work that way. There is no logically *necessary* connection between the typical policy platforms we see in most liberal democracies. For instance, we often observe that parties on the left of the political spectrum are in favour of both higher immigration and greater welfare spending. There is no necessary entailment of these positions. We can imagine voters who believe in all three of the possible other configurations of positions one might take with respect to these two issues. That is, one might be in favour of higher immigration and lower welfare spending (a typical libertarian view) or lower immigration and lower welfare spending (a typical conservative view) or lower immigration and higher welfare spending (a typical view among some communitarians and even some Rawlsians). If we can conceive of individuals with such views, then the mutual entailment assumption with respect to these two policy areas is unsupported and we cannot invoke it to rule out the possibility of voters trading off favoured against disfavoured policies.<sup>91</sup> The fact of someone voting for a pro-immigration, pro-welfare spending party is consistent with their favouring both policies but also consistent with the libertarian view so long as the libertarian valued expanding immigration more than restricting welfare spending or the communitarian view so long as the communitarian valued increasing welfare spending more than restricting immigration. What is true of these two policy issues seems true also of other typical policy correlations – e.g. policy programmes aiming to liberalise abortion *and* privatise state-owned enterprises or those aiming to increase taxation *and* adopt a more

---

<sup>89</sup> Though we still don't know the *relative strength* of their preferences for these policies.

<sup>90</sup> Perhaps policy B would amount to  $\neg A$ .

<sup>91</sup> Of course, that there is no logical contradiction between one's views on two issues doesn't imply that those views are necessarily *true*.

liberal approach to drug policy.<sup>92</sup> Therefore, the mutual entailment assumption appears to be without foundation. All this is not to deny that voting patterns reveal *something* about voter preferences and by extension reveal to liberal citizens some policy areas over which they disagree. We can, of course, safely infer from British and American voting patterns that UK citizens are divided over the issue of Brexit and US citizens over gun control. The point is that once we take into account the inevitable fact of voter compromise in a restricted political marketplace, our power to make such inferences is much less than often claimed.

One suspects that political philosophers who place ethical disagreement at the centre stage of political life simply generalise to society-at-large the experience of intellectual life within the academy. Here we have a small subset of the wider population united by a shared vocation, shared standards of inquiry and sustained social interaction in fora like academic conferences. They specialise in forming comprehensive perspectives over complex and difficult subject matter, the sort over which good-faith reasoners will arrive at divergent outcomes. As specialists in the business of arguing, they also spend a great deal of time debating with one another to try and establish who is right, though more often than not this debate leads to the entrenchment of rival positions rather than the conclusive victory of one view. It is under these specific conditions that we can identify a community of interlocutors with shared knowledge who experience disagreements in the relevant intersubjective sense. Indeed, given the specialised nature of the academy, leading academics often know their interlocutors personally. However, academic disagreements are to a large extent self-contained in the particular community in question. To take the example of political theory, the intricacies of the distinctions between left-libertarianism and Rawlsianism, classical liberalism and right-libertarianism or luck- and deontic-egalitarianism typically make sense only to the participants within the discipline. Members of the public-at-large are not party to such disagreements. Indeed, neither are many academics who specialise in different

---

<sup>92</sup> I must confess I struggle to think of any policy propositions or even more abstract political philosophy propositions that strictly entail one another. Even those that seem to do so initially are subject to counter-examples. For example, one might think that the propositions that “a property rights regime leads to much worse consequences for human welfare than socialism” and “we ought to replace capitalism with socialism” entail one another, such that someone who believes one of these would by that fact have reason to believe the other. But one can conceivably affirm the first and reject the second if they think, their terrible *consequences* aside, private property rights are nonetheless still *deontologically* justified – I owe this observation to Jason Brennan (2014b). Alternatively and perhaps less outlandishly, one could affirm the first but also think that, say, we are unable to implement socialism and that the best system available to us currently is capitalism. As Wittgenstein (1958) observed, our beliefs are connected with one another through webs of judgements rather than tight logical entailment.

fields. The epistemic relationship of a specialised community of academic inquirers maps quite poorly unto the relationship of liberal citizens *en masse*.

#### 4. Alienation Under the Emergent Account

We will now outline the emergent understanding of political society to be derived from Hayek's work. Mutual ignorance rather than intellectual disagreement is the primary source of alienation on this account. As a precursor to our discussion of Hayek, it would help to begin with a thinker with whom he has some perhaps unexpected similarities, Iris Marion Young.<sup>93</sup> She and Hayek disagree of course over economic policy and indeed over the status of liberalism itself. However, in *Justice and the Politics of Difference* Young makes some methodological claims about the way political and social theory is conducted that are strikingly similar to those of Hayek. She criticises political theorists, both liberal and communitarian, for pursuing what Michel Foucault called the Rousseauist dream, citing the following passage of the former:

It was the dream that each individual, whatever position he occupied, might be able to see the whole of society, that men's hearts should communicate, their vision be unobstructed by obstacles, and that the opinion of all reign over each.<sup>94</sup>

For Young, the Rousseauist dream is utopian in the pejorative sense because it assumes a "co-presence of subjects" which is possible at most only in face-to-face societies.<sup>95</sup> It fails to account for the sociological fact that members of impersonal social orders are opaque to one another. Against this, she argues that "politics must be conceived as a relationship of strangers who do not understand one another in a subjective and immediate sense, relating across time and distance".<sup>96</sup> Time and space, in her parlance, *mediate* the relationship between persons, rendering them strangers to one another. By the same token, persons enjoy unmediated relations with one another in face-to-face settings where they *do* have a "subjective and immediate sense" of one another. Now, Young recognises that *totally* unmediated relations are impossible - "[even] a face-to-face relation between two people is mediated by voice and gesture, spacing and temporality".<sup>97</sup> We can never see things *literally* from another's first-personal perspective and their account of it to us, no matter how rich or

---

<sup>93</sup> See Tebble (2003) for an extended comparison of the Young and Hayek, upon which this section draws. This section has profited greatly from Tebble's discussion there.

<sup>94</sup> Cited in Young (1990, 229).

<sup>95</sup> (Young 1990, 231)

<sup>96</sup> (Young 1990, 234)

<sup>97</sup> (Young 1990, 231)

immediate, will always remain a representation. Her point, I think, is that we can distinguish relationships by the degree of mediation they entail. While we each have unfiltered access to the soul of no one, perhaps not even our own, this doesn't prevent us making a firm distinction between the degree of mediation and transparency available between those who enjoy the spatio-temporal proximity of face-to-face relations and those who do not. In this sense, we can say that relations between members of impersonal orders are spatio-temporally mediated in a manner that those occurring between members of face-to-face orders are not.

On the basis of this epistemic distinction, Young accuses political philosophers of a modelling error. They assume members of large-scale orders have mutual knowledge of a magnitude apropos of smaller-scale settings. Once we realise this, we need to model differently such large-scale orders. Rather than romanticising the idyllic *Gemeinschaft* or the deliberative *polis*, she suggests that we look to the contemporary city as a normative ideal.<sup>98</sup> She does not intend by this a normative defence of sociological urbanisation or city life *as such*. Rather, I take her to see social relations within the city as a valuable metaphor for how political and social theorists should model large-scale impersonal orders that will, of course, contain within them both urban and rural spaces. She defines city life as the "being together of strangers".<sup>99</sup> In cities a very wide array of people from different social backgrounds, communities and groups interact with one another as part of very complex schemes of social cooperation without any deep sense of social unity or of participation in a common project:

City life is a vast, even infinite, economic network of production, distribution, transportation, exchange, communication, service provision, and amusement. City dwellers depend on the mediation of thousands of other people and vast organisational resources in order to accomplish their ends. City dwellers are thus together, bound to one another, in what should and sometimes is a single polity. Their being together entails some common problems and common interests, but they do not create a community of shared final ends, of mutual identification and reciprocity.<sup>100</sup>

Political theorists often model public spaces as a "realm of unity and mutual understanding". Young notes that this is not our experiences of public spaces in urban settings in which one encounters a diversity of opaque others.<sup>101</sup> Indeed, it is the possibility of encountering unknown or new experiences that lends to city life its "eroticism".<sup>102</sup> The normative ideal of

---

<sup>98</sup> (Young 1990, 238)

<sup>99</sup> (Young 1990, 237)

<sup>100</sup> (Young 1990, 238)

<sup>101</sup> (Young 1990, 240)

<sup>102</sup> (Young 1990, 239)

the city doesn't posit citizens engaged in any kind of common political project nor even trying to overcome their ethical differences so as to embark on such a project. It posits the social cooperation of strangers who know little of one another. This is a more appropriate methodological starting point, she thinks, for thinking about impersonal orders than the assumption of co-presence found in both communitarian and left-liberal political theory.

Hayek's understanding of an extended liberal order has much in common with Young's normative ideal of the city. His central inquiry, as expounded in works such as *The Constitution of Liberty* and *Law, Legislation and Liberty*, is to understand how coordination takes place among a mass of dispersed and spatio-temporally separated individuals, each of whom acts on their own localised knowledge. He understands a liberal order as an extended network of exchange and interaction between countless strangers too ignorant of one another to constitute, in Young's words, "a community of shared final ends, of mutual identification and reciprocity". He famously argued that socialist central planning was infeasible because the knowledge that the planning agency would need in order to formulate and execute a comprehensive economic plan was unavoidably dispersed and fragmented. It is not possible for a sovereign actor – be it the Politburo or the demos as a united body – to aggregate this knowledge and act upon the totality of it. I am going to term this Hayek's *fragmentation of knowledge* thesis. What Hayek has in mind here is not, for the most part, the fact that theoretical knowledge is unavoidably dispersed as a result of academic specialisation. He thought that under modernity we tend to be overly bewitched by this kind of knowledge, exaggerating its importance in human affairs.<sup>103</sup> The referent of the fragmentation of knowledge thesis is in many ways more prosaic and practical forms of knowing that can be understood in contradistinction to those forms expressible in explicit and general principles.

Such knowledge is of "the particular circumstances of time and place". In the economic context it means to "know of and put to use a machine not fully employed, or somebody's skill which could be better utilized, or to be aware of a surplus stock which can be drawn upon during an interruption of supplies".<sup>104</sup> Knowledge of such particulars are readily enjoyed by actors who are, as it were, *in-situ*, Hayek referring to them as the "man-on-the-spot." In respect of this kind of knowledge, "practically every individual has some advantage over all others in that he possesses unique information of which beneficial use might be

---

<sup>103</sup> (Hayek 1945, 521); (Hayek 2006 [1960], 24); c.f. (Oakeshott 1991)

<sup>104</sup> (Hayek 1945, 522)

made”.<sup>105</sup> Much of it is also tacit. In terms of Gilbert Ryle’s famous distinction, it is know-*how* rather than know-*that*.<sup>106</sup> Holders of such knowledge might struggle to verbally communicate it to outsiders. It is embodied in organisational routines and workplace practices, as well as in traditions, customs and conventions more generally. In successfully acting in the world we are as much reliant on such knowledge as embodied in “[our] habits and skills, our emotional attitudes, our tools, and our institutions” as we are on our theoretical knowledge.<sup>107</sup> Importantly, such knowledge is often of fleeting and temporary relevance. The knowledge of *surplus* stock, for instance, might be useful only *during* an interruption to the supply chain. Much of it is also “frequently contradictory”.<sup>108</sup> Facts of the world are never merely given but require interpretation. The interpretations that actors adopt depend much on the context and what they are trying to achieve. The same plot of land can be a holy site for pilgrims, an investment opportunity for a real estate agent and something in need of conservation for an environmentalist.

One may be tempted to see these forms of knowing as primarily instrumental. We use them as a means to realise our higher ends and values, which can themselves be couched in general principles, much like a Rawlsian comprehensive doctrine. Hayek thought this mistaken. Much of our understanding of the good life or moral conduct is tacitly embedded in the forms of life in which we participate. As already mentioned, very few of us subscribe to a Rawlsian comprehensive doctrine covering “all recognized values and virtues within one precisely articulated system”.<sup>109</sup> While we certainly have a practical understanding of what gives our life value and the moral goods we wish to realise, we may struggle to explicate in the form of a doctrine all such commitments. Hayek observes that in our ordinary life we “are constantly choosing between different values without a social code prescribing how we ought to choose”.<sup>110</sup> In any case, even those who try to live their life according to abstract principles have to interpret them in concrete settings which in turn requires knowledge of particulars. The substance of their conception of the good will be practical knowledge. One doesn’t just follow the priestly vocation in the abstract, one does so in a particular parish with a particular group of parishioners in a particular location and so on.

---

<sup>105</sup> (Hayek 1945, 522); c.f. (Hayek 2006 [1960], 26–27)

<sup>106</sup> (Ryle 1945)

<sup>107</sup> (Hayek 2006 [1960], 24)

<sup>108</sup> (Hayek 1945, 519)

<sup>109</sup> (Rawls 1996, 13)

<sup>110</sup> (Hayek 2001 [1944], 60–61)

Such knowledge is not inevitably fragmented amongst members of a group. Size and the degree of heterogeneity matter. In small groups where members are not spatio-temporally distant, share the same local circumstances and social practices, and where transaction costs are low, practical knowledge can be relevantly shared. This is so even with respect to its tacit component, which can be transmitted and learnt via imitation. Eugen Weber claims that members of the typical nineteenth century French village barely needed to speak with one another, such was their intuitive understanding of one another's behaviour:

[T]he peasant shows what he does or is about to do in forms of behaviour that are wholly familiar to his fellows and so are easily interpreted by all of them. If one sees a man in such a place, at such and such a time, one can generally tell what he is up to. Deduction based on concrete observation is what counts, and speech adds little to this.<sup>111</sup>

By the same token, however, this epistemic transparency is lost if the group in question are the impersonally related members of contemporary liberal societies. Few would deny that the fragmentation of knowledge under such circumstances poses difficulties for large-scale social coordination. Many think, however, that the relevant knowledge can be aggregated by the use of two potential mechanisms. The first is effective statistics. In aggregating information, such mechanisms overcome the fragmentation of knowledge. However, it is precisely the knowledge of particular circumstances for which standardising statistics struggle to account. Hayek writes:

The statistics which such a central authority would have to use would have to be arrived at precisely by abstracting from minor differences between the things, by lumping together, as resources of one kind, items which differ as regards location, quality, and other particulars, in a way which may be very significant for the specific decision [that they are looking to undertake].<sup>112</sup>

Additionally, given its temporary relevance much pertinent knowledge would lose its relevance once it reaches the central authority and the tacit component is not even amenable to central aggregation.<sup>113</sup> The second is democracy. We shall have much more to say about this in the chapters to follow, but for now we can make the following observations. To the extent that the fragmentation of knowledge refers to tacit knowledge, it is hard to

---

<sup>111</sup> (Weber 2007, 92)

<sup>112</sup> (F. A. Hayek 1945, 524); by analogy we argued above that the standardised nature of choices available in the ballot box are a poor proxy for each person's concrete and possibly even idiosyncratic political preferences.

<sup>113</sup> Recalling the old military dictum that the generals are always prepared to fight the previous war. For an extended analysis of the blind-spots of synoptic state rationality, particularly with regard to local and practical knowledge, see Scott (1998).

see how it can be conveyed effectively through voting or deliberation.<sup>114</sup> To the extent that it is fleeting knowledge of particular circumstances, it is hard to see how each citizen can convey it to all her fellow citizens in a timely fashion. To the extent that it is perspectival knowledge dependent on what one is trying to achieve, it is hard to see how such knowledge can be aggregated unless one particular perspective is privileged. Furthermore, deliberative democrats such as Jürgen Habermas and Joshua Cohen concede that democratic citizens can properly pool their knowledge to form a genuine collective will only in face-to-face groupings where they are comparatively epistemically transparent to one another.<sup>115</sup>

We will be at pains to show in this thesis that the fragmentation of knowledge would prevent liberal citizens acting in unison even if they all individually affirm the same principles of justice or some other normative value. For there would still remain the problem of signalling this fact to one another.<sup>116</sup> The arguments from section two above have already cast doubt on the possibility of this happening. Chapter four will develop these claims. Further, even if this problem were overcome and liberal citizens *did* become mutually aware of their abstract commitment to justice, they would still face a knowledge problem in *coordinating* their actions to attain that end.<sup>117</sup> It is of great significance that the emergent view sees mutual ignorance rather than disagreement as the primary barrier to collective political agency. It means that the emergent view is much less salutary about our prospects for attaining it. If disagreement is the primary issue, then the problem seems at least in principle soluble. We just find principles on which we can agree. If, however, what prevents us from acting *en masse* on shared intentions is mutual ignorance, the prospects of a solution seem even less promising. The tacit, fleeting and localised nature of the practical knowledge held by each citizen precludes it from being communicated to all the others. Our inability to take collective control is a fact of life with which we would need to make do. Or so the analysis to follow will argue.

---

<sup>114</sup> This claim is elaborated by Pennington (2003).

<sup>115</sup> I offer an extended critique of attempts to scale-up deliberative democracy in chapter four of this thesis.

<sup>116</sup> Per our analysis from section two, that they individual affirm the same principles does not mean that they agree in the relevant intersubjective sense.

<sup>117</sup> (Tebble 2016b, 19)



## 5. The Emergent Account and the Function of Justice

On the associational understanding, we turn to abstract and impartial principles of justice to overcome intellectual disagreement. As goes the familiar argument, a self-governing political association cannot be based on the principles of utilitarianism or Catholicism or even too robust a commitment to autonomy since each of these will be reasonably rejected by some sincere citizens. Therefore, we abstract away from our particular disagreements to find governing principles that are suitably impartial and to which all citizens can assent. The currency of Rawlsian justice, for instance, are primary goods that purportedly anyone would want whatever else they want. On the emergent understanding, the rationale of abstract principles of justice are rather different. The problem is not so much that I disagree with those I cooperate with but that I know little about them. We typically interact with strangers in a liberal society. Far more often than not, we don't know the conception of the good or religion of the person sitting next to us on a train or in front us in a queue or who sells us something in a shop. We often don't know this information even of our co-workers and neighbours. Here I refer only to those with whom we consciously interact. In an economy characterised by an even moderate division of labour, we have mediated relations with countless others apart from those with whom we have direct relations. Of these unseen others we often have scarcely *any* knowledge. Our interactions with all these strangers are facilitated by abstract and impartial rules of justice. It does not seem right, however, to say that justice resolves disagreement. After all, if they are strangers we typically don't know their views on ethics or other matters.

These observations make clear that the function of justice under the emergent understanding is to facilitate interaction between strangers despite their mutual ignorance. The puzzle to be solved is this. If each agent acts on private knowledge unknown to all other agents save those in her immediate vicinity, then one might wonder how agents can coordinate their individual plans so as to be compatible with one another. Hayek spends his whole career trying to show coordination under such conditions is, in fact, possible so long as agents exhibit some regularity of behaviour in the manner in which they act upon their otherwise idiosyncratic circumstances and knowledge. It is possible, in other words, so long as the agents subscribe to common *rules* when acting on their localised information.<sup>118</sup> If such rules are to be understood and followed by numerous agents acting under variegated

---

<sup>118</sup> (Hayek 1982a, 39–44)

circumstances, these rules of *just conduct* will have to display a certain purpose-independence.<sup>119</sup>

A helpful metaphor once used by him was the highway-code.<sup>120</sup> Any road network will have very many road users with very many different destinations. No driver knows nor do they need to know the destination of the other drivers. Nor is there a global controller who knows the destination of each of the drivers. Still, the vast majority of road journeys end at their destination. Road users successfully coordinate with one another because they subscribe to common rules of the road e.g. they all drive on the left- or right-side. Such rules enable countless road users to generate reasonable expectations of one another's behaviour and successfully coordinate their actions despite their mutual ignorance of one another's destinations. They are purpose-independent rules insofar as they do not mandate any particular destination but specify constraints on how one arrives at any destination. *Mutatis mutandis* in a liberal social order, adherence to common rules allows individuals and communities to successfully interact and coordinate with one another while acting on their own idiosyncratic knowledge and values. Hayek's understands these "rules of just conduct" in a fundamentally Humean manner. They relate to the stability of possession, the transfer of property by consent and the fulfilment of contractual promises. They are abstract and instrumental rules that do not specify ends to be pursued but conditions to be adhered to in such pursuit. In this regard they facilitate cooperation between agents from a wide variety of ethical and cultural frameworks. Often they are best understood as negative prohibitions. Rather than *prescribing* actions-to-be-performed, they merely *proscribe* a limited range of actions – e.g. stealing or trespassing – while otherwise granting individuals freedom of action. There are three ways that conformity to such rules facilitates coordination between epistemically alienated liberal agents.

a) Economising on interpretive labour in large-scale settings

I borrow the term "interpretive labour" from David Graeber.<sup>121</sup> Social life is possible only if we can understand the meaning and intent behind one another's actions. We need to be able to see things from the perspective of another person and discern what they meant in performing a particular action or making a particular utterance if we are to respond

---

<sup>119</sup> (Hayek 1982a, 80–81). Hayek long thought that the propensity to follow abstract rules was deeply embedded in human nature. In his work on theoretical psychology, *The Sensory Order* (1987 [1952]) and his paper "rules, perception and intelligibility" (1967a), Hayek advanced the claim that the mind itself is a device that classifies external stimuli according to abstract rules of classification.

<sup>120</sup> (Hayek 2001 [1944], 78); c.f. (Schmidtz 2010, 84–86)

<sup>121</sup> (Graeber 2015, 41); c.f. (Rawls 1999, 409–10)

appropriately. Interpretive labour is the process or act of gaining such knowledge of another actor. Social life requires not just that we engage in such labour but that we do so *efficiently*. In economic terms, we can view interpretive labour as a transaction cost which needs to be minimised if we are to secure the benefits of social cooperation. Predictable and conventional norms of behaviour fulfil this economising function. When someone I have just met initiates a hand shake, they are signalling by such a customary act their intention to greet me. Not only does their customary behaviour efficiently signal their intention to me but also readily available is an act-type on my part that efficiently signals to them my intention to reciprocate their greeting. I simply raise my own hand to meet theirs. Customary rules perform this economising function only if they are affirmed and understood by those party to the interaction. The gesture to shake someone's hand could not effectively signal intent if directed to someone unfamiliar with the custom. That we behave in *such-and-such way* in *this* kind of situation needs to be common knowledge between us.

In small-scale societies the relevant rules can be derived from the shared cultural and religious framework of the group, as anthropologists have long argued. According to Marcel Mauss, for the Native American Tribes of the American Northwest and the Polynesian tribes of Papa New Guinea, economic exchange was "only one element" of interactions that extended to "banquets, rituals . . . dances, festivals, and fairs".<sup>122</sup> These elaborate rituals allow members of these groups to reliably signalling their intent to one another and thereby coordinate their actions. The drawback here is that *intra*-group exchange and interaction is made difficult since an outsider will lack knowledge of the group's customs and rituals. A foreigner would, for example, find it exceedingly difficult to join the much discussed gift economy of the Kula Ring of the Tobriand islands in Papa New Guinea. It consisted of a highly complex barter economy the rules of which were derived from the elaborate and particular cultural framework of the several thousand inhabitants of those islands.<sup>123</sup>

Efficient rules of interpretive labour in a large-scale and diverse society must take a different form. They need to be understood and easily subscribed to by followers of a heterogeneous and variegated array of cultural forms. They cannot be embedded within any particular cultural form but should exhibit a certain generality. Hand-shaking is a good example of such a rule. It can be easily learnt and internalised by people from many walks of life. I might know very little about the person who initiates a hand-shake. Still I instantly recognise what they

---

<sup>122</sup> (Mauss 2002, 6–7)

<sup>123</sup> (Mauss 2002, chap. 1)

are trying to do. Analogously, the same can be said of generalised rules to do with possession, exchange and promise-keeping. They constitute norms that allow diverse individuals to successfully signal their intentions to one another. I do not know about the religion, hobbies, interests or values of the Uber driver whose services I procure and vice-versa. I could even be a tourist residing in the country temporarily. Nonetheless subscription to generalised rules of exchange allows us, at least most of the time, to successfully interact with one another. This is the great and often overlooked benefit of the cash nexus.<sup>124</sup>

b) Reducing the need for collective decision-making

Hayek writes of the importance of “general rules governing the conditions under which objects or circumstances become part of the protected sphere of a person or persons”. He refers, of course, to rules demarcating private property rights. When such rules are accepted in the community, they enable “each member of a society to shape the content of his protected sphere and all members recognize what belongs to their sphere and what does not”.<sup>125</sup> Hayek’s claim is that given the inevitable fragmentation of knowledge, any complex and large-scale society would require property rights, whether these are formally or informally recognised.<sup>126</sup> The reason for this is that it is difficult to carry out a “coherent plan of action unless we are certain of our exclusive control of some material objects; and where we do not control them, it is necessary to know who does if we are to collaborate with them”.<sup>127</sup> In a similar vein, Gerald Gaus conceptualises property rights as the partitioning of the social space into different jurisdictions each of which is governed by a particular right-holder.<sup>128</sup> Jurisdictional rights are purpose-independent insofar as they provide each rights-holder with significant latitude as how to arrange their jurisdiction. Such partitioning helps us deal with the fragmentation of knowledge in large-scale societies by alleviating, at least somewhat, the need for the coordination of plans.

Supposing the property regime does a reasonable job of internalising the costs and benefits of the actions of each rights-holder, each can govern her jurisdiction as she wishes without worrying about the countless alternative uses that might be made of the resources under

---

<sup>124</sup> As stressed by Hayek (1982b) throughout *Law, Legislation and Liberty* but particularly chapter 10 of volume 2.

<sup>125</sup> (Hayek 2006 [1960], 123)

<sup>126</sup> (Hayek 1988, chap. 2); this was de facto true even of the Soviet Union, see Roberts (2002, 261).

<sup>127</sup> (Hayek 2006 [1960], 123); of course the point can be generalised beyond “material objects” to intangible property claims.

<sup>128</sup> (Gaus 2012a, 370-380); the right-holder need not be a singular individual, as demonstrated by Elinor Ostrom’s (1990) work on the commons.

her control by other members of the community. In direct contrast, under a system of collective ownership, each will have to ensure that their utilisation of a resource is compatible with the potential plans that all other agents might have for it. This may be possible in very small-scale settings, such as the hypothetical camping trip valorised by Gerald Cohen in his short book *Why not Socialism?* Cohen imagines a small group of egalitarian campers who treat their camping facilities as under collective control for the duration of their trip. The group have shared understandings as to who should be able to use which resources for the duration of the trip as well as who should perform which chores.<sup>129</sup> Direct and conscious coordination on how resources should be used are feasible given the epistemic transparency that holds between the group. However, as Cohen himself acknowledges later in the book, this is precisely what is missing between members of a large-scale society.<sup>130</sup> Note that this justification for property differs from that of the associational approach. On the associational view, we need property because liberal citizens would be unable to reconcile their competing claims as to how their common resources should be utilised. On the emergent view, we need property because liberal citizens are *unaware* of the claims each might make as to how such resources should be utilised.

I believe we can generalise this epistemic claim to liberties other than property rights.<sup>131</sup> Let us take some that are prominent within Rawls' account of the basic liberties. These include "freedom of speech and assembly, liberty of conscience and freedom of thought [and] freedom of the person".<sup>132</sup> Now, it is at least conceptually possible to imagine that rights such as these were held not individually but collectively. One could not just decide for oneself, say, what to do with one's body or where to live or how to worship and so on but would have to gain the assent of the whole community for each decision. This would be an egalitarian society insofar as each person would have an equal stake in the "common pool" of societal resources. Perhaps something approximating this is possible in a small-scale community. Benjamin Constant's "liberty of the ancients" arguably was not far off this system of social organisation.<sup>133</sup> However, it is immediately apparent that this would cause grave coordination problems if transposed to a large-scale society. Any exercises of a

---

<sup>129</sup> (Cohen 2009, chap. 1)

<sup>130</sup> (Cohen 2009, chap. 4)

<sup>131</sup> Though the point here is moot if one adopts the strident libertarian position that all rights are reducible to property rights. I take the classical liberal view that there are individual rights not reducible to property or jurisdictional rights. On this distinction, see John Tomasi's *Free Market Fairness* (2013, Chps 2 &4).

<sup>132</sup> (Rawls 1999, 53)

<sup>133</sup> "Thus among the ancients the individual, almost always sovereign in public affairs, was a slave in all his private relations". (Constant 1819)

significant liberty would need the assent of millions of others if we take the “community” in question to be members of a typical nation-state. In Richard Epstein’s apt words, such a system of collective ownership “builds a form of institutional paralysis into the very fabric of social lives”.<sup>134</sup> Privatising these kinds of decision-making, again, circumvents this mammoth collective coordination problem. Each can decide how to use their body or which speech-acts to engage in or how to worship the divine using only their own information without having to gain the consent of all other members of society. Again, the difficulty is not so much that we might disagree with the countless other competing claims that might be made on the use of our “bodily resources,” but that we cannot fathom what they might be.

c) Transmission of dispersed information

This is perhaps the coordinative function of rule-following with which Hayek is most commonly associated. In small-scale social orders where parties had mutual knowledge of one another’s needs and contributions, economic distribution and adaptation to changing circumstances could be the product of conscious choice and design. Recourse to such methods is precluded in large-scale social orders due to the fragmentation of knowledge. Nonetheless, indirect coordination is possible so long as market actors respect Humean rules of just conduct and exhibit the following regularity in their behaviour - “they will normally prefer a larger return from their efforts to a smaller one, and . . . they will [often] increase their efforts in a particular direction if the prospects of return improve”.<sup>135</sup> The unintended result of subscription to these common rules is the formation of *signals* in the form of market prices that reflect tacitly, and albeit imperfectly, the dispersed information of countless actors otherwise mutually ignorant of one another’s circumstances. Hayek gives the famous example of a change in the price of tin.<sup>136</sup> Suppose that some unknown event renders the supply of tin comparatively scarce. The price of tin rises. The price rise is passed down to the consumer through the successive profit-maximising behaviour of agents down the supply chain. The consumer of tinned goods doesn’t need to be aware of either this distant event or the perturbations it has caused in the tin and other related industries. They simply take note of the increase in price and economise their behaviour accordingly. In this way they incorporate in their behaviour information of which they have no direct apprehension.

---

<sup>134</sup> (Epstein 1995, 56)

<sup>135</sup> (Hayek 1982a, 45)

<sup>136</sup> (Hayek 1945, 526)

Wide spread subscription to liberal rules of just conduct do not transmit information merely through price signals. They also facilitate widespread social experimentation. We all potentially benefit from the emergence and adoption of new social practices that emerge from this process. This is the theme that Hayek stresses particularly in the first three chapters of *The Constitution of Liberty*. The freedom of action afforded by rules of just conduct allows a wide range of responses to social problems and challenges based on the local knowledge of agents *in situ*. This is particularly useful in aiding adaptation to changing circumstances: “[every] change in conditions will make necessary some change in the use of resources, in the direction and kind of human activities, in habits and practices”.<sup>137</sup> The experimentation afforded by rules of just conduct maximises the possibility of discovering new social practices to respond to changing circumstances. These can then be transmitted to society-at-large via processes of imitation. Crucially, the adopters need not know what makes this particular rule or norm successful nor of the alternative attempts that did not prove successful. They tacitly “gain” such knowledge simply by imitating the practice. Such processes occur not simply in the economic sphere but also in the moral sphere. One need only consider the rapid changes in, say, gender norms or attitudes to homosexuality that have occurred in many parts of the world in the last few decades. In recent years, Hayekian scholars have applied this insight unto questions of social justice, arguing that the Rawlsian ideal of a well-ordered society united by one conception of social justice is defective. In a complex and open society, our understanding of distributive justice should be open to change and evolution both to be able to respond adequately to changing circumstances and to take advantage of the benefits of social learning and experimentation.<sup>138</sup> Again, we see the epistemic benefits of the comparative purpose-independence of the rules of just conduct. We need rules open to a wide range of experimentation precisely because we do not know *a priori* which experiments will prove useful. If we did, there would be no need for the experiments in the first place.

I wish to end this section by noting an important distinction. One can accept Hayek’s argument that abstract rules of just conduct are needed to facilitate interaction under impersonal settings without assuming that all these rules should originate from the same source – e.g. the state. Thus, the rules governing transactions on an app-based taxi platform, stock listings on a stock exchange, the demarcation of property rights in a homeowners association or credit card transactions around the world might all be described as rules of just conduct in the Hayekian sense. They facilitate interactions between many agents and

---

<sup>137</sup> (Hayek 2006 [1960], 26)

<sup>138</sup> On this see Pennington (2011) and Gaus (2019).

groups with a wide range of different purposes. Yet none of these rules originate from the same source nor should they all be seen as the application of one general scheme of property rights. Hayek himself was not as clear on this as one would like.<sup>139</sup> The point here is to resist the temptation to connect the exigencies of governance in large-scale societies with a legal centralism that sees the state as the only important maker and enforcer of institutional rules. The importance of this distinction will become apparent in the following chapter, where we use Hayek's ideas to criticise the idea that there is a monocentric source of institutional rules in society.

## 6. The Emergent Account and Unintended Outcomes

The associational view sees political institutions as the product of the coordinated agency of liberal citizens who act upon and realise their shared intentions. The emergent view stresses that large-scale social outcomes are the result of human action but not intentional design. Hence, Hayek stresses that the outcomes or "order of actions" to which interactions under the rules of just conduct give rise are not reducible to the intentions of the parties whose actions are causally responsible for them. In this sense, these outcomes are emergent and unavoidably so. Trying to directly engineer a particular manifestation of the order of actions is impossible if not very difficult since such outcomes will "always be an adaptation to a large number of particular facts which will not be known in their totality to anyone".<sup>140</sup> Each acts on the basis of some knowledge unknown to the rest and this knowledge cannot be pooled in impersonal settings. We have already made note of how such pooling is precluded by the tacit, localised and fleeting nature of this information. Another important factor is that the data itself is constantly in flux. The countless agents-in-question are constantly revising and updating the information they act upon as a result of their continuous interactions with one another.

Now, Hayek is sometimes interpreted as saying that these emergent outcomes, or spontaneous orders to use his famous terminology, are black boxes about whose operation we can barely say anything.<sup>141</sup> This interpretation is not correct. Hayek does think that we

---

<sup>139</sup> For an excellent critique of Hayek's legal centralism, see Stringham (2015, chap. 13).

<sup>140</sup> (Hayek 1982a, 40)

<sup>141</sup> As already mentioned, emergent phenomena "exhibit[s] properties that are different from those displayed by the elements [of which they are composed] taken in isolation" (Lewis 2016, 115). Hayek himself rarely used such terminology, preferring his better known term of "spontaneous order". This thesis tends to adopt the "emergence" phrasing because it emphasises, in a way that I don't think the notion of "spontaneity" does, the *disjuncture* between the individual intentions of liberal agents and



can achieve an “explanation of principle” with regard to such phenomena.<sup>142</sup> We can understand the character of different institutional rules and say something about the broad range of outcomes or “orders of actions” to which they might give rise, without being able to predict any particular manifestation. Thus, we can make the broad based prediction that market-based institutions outperform institutions of central planning. However, and as attested to by the dire record of economists in this area, we cannot predict specific growth rates. In this respect, Hayek’s social theory is widely regarded as anticipating the science of complex systems that is increasingly influential in fields ranging from computer science to cognitive psychology.<sup>143</sup> Scholars in this field have formally shown that complex systems composed of interdependent variables are compatible with a very large range of conceivable outcomes, limiting greatly our ability to predict their outcomes.<sup>144</sup>

On this Hayekian view, the aspiration for collective political agency and social unity is based on an exaggeration of our ability to predict let alone control the outcomes of large-scale coordination between epistemically distant strangers. Particularly in *Law, Legislation and Liberty*, Hayek tries to offer a different understanding of a liberal order. It consists of agents whose political unity does not extend beyond common subscription to instrumental rules of just conduct. This order constitutes a “Great Society” because under modernity there is no “single” society to which we belong but most of us are members of “many different overlapping and interlacing societies.”<sup>145</sup> The Great Society, then, will consist of many groups with sufficient mutual knowledge of one another’s purposes and circumstances such that their internal governing rules would promote substantively shared ends. The order as a whole, however, is more akin to a decent system of traffic lights than an association in pursuit of shared ends. Iris Marion Young’s proposed re-orientation of political theory was already to be found in Hayek’s *Law, Legislation and Liberty*. For the Great Society is a model of political society that does away with the assumption of co-present subjects.

At this juncture it is important to also note that Hayek’s thinking about the extent to which the rules of just conduct can be rationally designed changes during his career. The early Hayek of *The Road to Serfdom* and of *The Constitution of Liberty* is much friendlier to rational

---

the outcomes to which their interactions give rise. In this way the distinction between the associational conception is rendered more vivid.

<sup>142</sup> (Hayek 1955); (Hayek 1967b 29-31)

<sup>143</sup> On this see Vaughn (1999) and Lewis (2016, 2012).

<sup>144</sup> “A mere 10 variable equation system can generate approximately 181,440 possible solutions.” (Vaughn 1999, 242 (n4))

<sup>145</sup> (Hayek 1982c, 476)

design of such rules in order to make maximally beneficial use of spontaneous ordering forces. On this early Hayekian view, “explanations of principle” can help us predict, albeit in broad-based terms, the order of actions arising under different rules of just conduct. They provide us with a rational basis to socially engineer such rules to obtain outcomes we consider more desirable. Though such engineering will always fall short of planning for any *specific* order of actions. Hence, he speaks of “planning for competition” in *the Road to Serfdom* and dedicates a third of *The Constitution of Liberty* to providing quite programmatic and detailed policy prescriptions to legislators.<sup>146</sup> There is a distinct change of emphasis, however, by the time Hayek writes *Law, Legislation and Liberty*. There he conceives of the rules of just conduct as evolving slowly over time through the countless case-by-case decisions of common law judges. In other words, the later Hayek begins to think that both the order of actions *and* the underlying rules of just conduct are the emergent result of many human actions but no overall design. This is not to say that he completely eschews legislation in his later work. However, the task of legislation is now the more limited one of “correcting” “grown” law when it does not adapt to new circumstances fast enough or begins to develop in undesirable directions.<sup>147</sup> With this distinction in mind, the argument of this chapter might be subject to the following objection. A reconciliation between Hayekian social theory and the Rawlsian associational model might be possible if we adopt the early Hayekian view. We might then conceive of liberal citizens as enjoined in a political association under which they rationally and collectively decide the rules of just conduct that would govern their interactions in the otherwise spontaneous order of civil society. This, roughly put, is how I believe we can interpret John Tomasi’s work, *Free Market Fairness*.<sup>148</sup> Chapter two is dedicated to rebutting this important objection. It defends a position akin to that of the later Hayek. It argues that the system of institutional rules governing a liberal order, and not simply the order of actions to which such rules give rise, should be seen as emergent phenomena rather than the rational construct of a civically-engaged liberal citizenry.

## 7. Hayek’s Worry

Hayek does not himself use the term “collective political agency” but I think it captures well the kind of politics that he thought was unfeasible in the contemporary world, namely one

---

<sup>146</sup> (Hayek 2001 [1944], 43); (Hayek 2006 [1960], part. 3)

<sup>147</sup> (Hayek 1982a, 88–89); Hayek also stresses the emergent and evolved nature of the rules of just conduct in his final work, *The Fatal Conceit* (Hayek 1988, chap. 3).

<sup>148</sup> (Tomasi 2013)

that aspires to subject large-scale and unavoidably emergent social processes to some form of collective control. The difficulty with this kind of politics for Hayek is not simply its unattainability. The issue is that the attempt to realise it would undermine a free society. When Hayek writes *The Road to Serfdom*, the desire for collective political agency is reflected in the demand for socialism. That work is famously dedicated to showing the threat to freedom stemming from such politics. In the latter half of the twentieth century, when the failures of command-and-control socialism were becoming all too apparent, collective political agency was espoused to realise social justice. Hence, *Law, Legislation and Liberty* is dedicated to showing the threat to freedom posed by patterned conceptions of social justice. While for Hayek the goals of such politics will always prove to be chimerical, the restrictions on individual freedom that arise in the pursuit of such goals will be all-too real. Both in *Law, Legislation and Liberty* and in his final work *The Fatal Conceit*, Hayek also begins to evince scepticism about our ability to resist the lure of the politics of collective control.

By this late stage of his career, Hayek thought that the desire for collective political agency is an evolutionary inheritance. Our core moral intuitions are adapted for life in the face-to-face tribal groupings in which we evolved and in which unified action towards the “the direct pursuit of felt needs or perceived objects” was genuinely possible.<sup>149</sup> Members of liberal orders would find it difficult to discipline their evolutionarily atavistic moral intuitions when they are engaged in politics.<sup>150</sup> They would consistently be tempted to apply the “unmodified, uncurbed, rules of the micro-cosmos (i.e., of the small band or troop, or of, say, our families) to the macro-cosmos (our wider civilisation)”.<sup>151</sup> If Hayek’s psychological thesis is correct, two important implications follow. First, liberal citizens will find it difficult to endorse the liberal order. Its workings will appear to them as something of an alien power that they cannot control. Second, they will be perennially attracted to political projects incompatible with liberalism. The institutional stability of a liberal order will be put into doubt. Since this thesis adopts Hayek’s stance that collective political agency is not possible

---

<sup>149</sup> (Hayek 1982c, 160)

<sup>150</sup> While this evolutionary conjecture strikes me as at least plausible, Hayek’s claim that the desire for collective political agency is deeply engrained in human nature need not depend on it. That the following is true seems sufficient: “human emotions are attached to concrete objects, and the emotions of justice in particular are still very much connected with the visible needs of the group to which each person belongs” (Hayek 1982b, 143). Indeed, the concern that the abstract relations of a liberal order are alienating is also found in figures such as Hegel, Marx and Weber. It is unsurprising that Hayek, who came of intellectual age in Vienna, would share this concern. Andrew Gamble (2007, 44-46 & 180-2) brings out well the parallels between Weber and Hayek.

<sup>151</sup> (Hayek 1988, 14); Gerald Cohen’s (2009) hope that we can model a large-scale society on the norms of a weekend camping trip is an interesting philosophical expression of this desire.

between impersonally related liberal citizens, it behoves us to engage with Hayek's concerns on this issue. Chapter seven will alleviate Hayek's concerns about the stability of liberal institutions. The conclusion to this thesis will provide some reflections on how we can personally reconcile ourselves to a social world which exceeds our collective capacity to control and direct it.

## **Conclusion**

A summary of the claims of a long chapter are in order. The chapter began by outlining the associational understanding of political society. This popular conception is found in the work of John Rawls and other prominent political theorists. It conceives of political society as an association through which citizens exercise collective autonomy over the basic structure of their social world. The primary barrier to achieving this great social good is cognitive disagreement. This is overcome by purpose-independent principles of justice that transcend the partisan intellectual differences of liberal citizens. In contradistinction to this view, the chapter then proposed the alternative, emergent understanding of political society rooted in the thought of Hayek. On this view, liberal agents are alienated primarily by mutual ignorance. Purpose-independent rules of justice are akin to traffic management systems that facilitate coordination between countless agents whose ends and purposes are opaque to one another. Political society cannot be seen as an association through which citizens exercise collective agency. For the fragmentation of knowledge between liberal agents precludes their acting on shared intentions. Instead, the outcomes of their interactions are unintended phenomena. The difficult upshot of the emergent view is our having to come to terms with this fact. Chapters 3-6 will attempt to vindicate the emergent understanding of political society over the associational one. However, before moving on to that task, chapter two will explain why this thesis has not followed those classical liberal philosophers who attempt to reconcile Hayek and Rawls.

I wish to end this chapter by making some remarks as to the wider, quite deflationary implications that the emergent view of political society poses not just for the notion of collective political agency but also for ideal theory, the approach to conducting political philosophy espoused by Rawls and others. These remarks will help also to distinguish the emergent view of political society from other Hayekian-inspired critiques of Rawlsian political methodology. An "ideal" theory of justice is supposed to be an inspiring and

demanding vision of the just society that can aid political reform in an action-guiding manner. The term was coined by Rawls in *Theory*.<sup>152</sup> His well-ordered society is supposed to be just such an inspiring vision.<sup>153</sup>

A theorist can idealise on various dimensions. As we have already seen Rawls idealises along the dimension of compliance. He imagines a well-ordered society in which all fully comply with its directives.<sup>154</sup> Laura Valentini helpfully identifies two further dimensions in which a normative political theory might be idealised.<sup>155</sup> One dimension of idealisation refers to whether the theory aims to identify a perfectly just “end-state” as opposed, more modestly, to instances of concrete *injustice* requiring reform. The latter alternative to end-state theorising is helpfully sketched out by Amartya Sen. He famously argues that Rawlsian end-state theory is neither necessary to combat concrete instances of injustice nor sufficient since the former’s distance from day-to-day political realities means it provides little or no action-guiding help in combatting such injustices.<sup>156</sup> The other dimension is the degree to which the *content* of the theory of justice incorporates real-world “factual constraints”. The more the theory’s prescriptions make allowances for factual and contextual constraints, the more non-ideal it would be on this dimension. From this perspective Rawls’ theory of justice is actually less idealised than that of a figure like Gerald Cohen. While the latter insists that the content of justice must be “fact-independent,” Rawls does allow certain factual constraints to determine the content of his two principles of justice.<sup>157</sup> One such constraint being the fact of moderate scarcity.<sup>158</sup> The assumption being that a theory of justice premised on the absence of scarcity – Marx’s second stage of socialism for instance – would not be practically useful for the kind of world that we inhabit.

Now, much of the criticism of ideal theory, particularly in its Rawlsian variant, has focused on the implausibility of full compliance. It is often argued that the pervasiveness of ethical disagreement as well as sheer self-interest in political life renders implausible any theory of justice that assumes their absence. As previously mentioned, the emergent view defended in this thesis assumes away compliance problems to highlight the epistemic shortcomings of

---

<sup>152</sup> (Rawls 1999, 7–8). While Rawls coined the term, he quite obviously did not innovate the approach. Philosophers and social thinkers have concerned themselves with outlining inspiring visions of normatively perfect societies since at least Plato.

<sup>153</sup> Thus we might say that for Rawls, the implementation of an ideal theory of justice is the *object* of collective political agency.

<sup>154</sup> (Rawls 1999, 215–16; 7–8) c.f. (Rawls 2001, 66)

<sup>155</sup> (Valentini 2012)

<sup>156</sup> (Sen 2009, *passim*)

<sup>157</sup> (Cohen 2003)

<sup>158</sup> (Rawls 1999, 110/226)

Rawls' approach to political philosophy. In this regard it is similar to other Hayekian-inspired critiques of ideal theory prevalent in the literature. The nub of these critiques, however, similar to that of Sen, have been on end-state idealisation. They point out that the theorist cannot know what the perfectly just society would like. End-state theories of justice beg the epistemic question posed by Hayek's thesis as to the fragmentation of knowledge in a large-scale society. Just as no one can aggregate the necessary knowledge to plan the production and distribution of goods and services in a complex economy, no one can aggregate the necessary knowledge to furnish a blueprint of perfectly just institutions. What we require are classical liberal meta-institutions which in utilising dispersed knowledge can help us *discover* which principles or institutions are the most just.<sup>159</sup>

The emergent view of political society is not in tension with these other Hayekian-inspired critiques. However, it takes a different tact. On the emergent view, a *further* epistemic barrier to attaining ideal theory would remain even if we knew for sure the correct end-state theory of perfect justice. The remaining barrier would be the inability of citizens to coordinate their actions to attain this given end due to the fragmentation of knowledge in a large-scale society. In terms of Valentini's typology, the emergent view criticises ideal theory not so much for positing an end-state of justice but for overlooking epistemic, factual constraints on the ability of citizens to harmonise their intentions and actions. Quite separate from the difficulty of identifying what the given ends of political reform should be, the problem on the emergent view is that there is no collective, identifiable *agent* for whom such a theory could be action-guiding. There is no practicable *duty-bearer* for ideal theory. This last claim will be elaborated upon in greater detail in the next chapter.

---

<sup>159</sup> The aforementioned works of Gaus (2019) and Pennington (2011) provide versions of this critique, as do Adam Tebble's account of "epistemic liberalism" (2016a) and Nozick's (2012 [1974]) famous meta-utopia in the last third of *Anarchy, State and Utopia*.

## Chapter Two: “Rawlsekianism” – a Critique

### Abstract

The prior chapter attempted to drive a wedge between the Hayekian and Rawlsian understandings of political society. The emergent understanding derived from Hayek’s thought emphasises the undesigned and spontaneous nature of political institutions. The associational understanding exemplified by Rawls emphasises that just political institutions stem from the collective political agency of liberal citizens. Some contemporary classical liberals have adopted a more conciliatory approach, combining Hayekian insights about emergent social phenomena with Rawlsian justificatory ambitions. This chapter considers and critiques two such attempts at a reconciliatory approach, those offered by John Tomasi and Gerald Gaus. The central claim is that institutional systems, like the order of interactions to which they give rise, are emergent phenomena. They cannot be made to accord with the shared intentions of the citizenry.

### Introduction

“Rawlsekianism” was a term coined by Brink Lindsey in an op-ed for the *New Republic* magazine on December 11 2006. At the time, Lindsey was vice president for research at the classical liberal US think tank, The Cato Institute. Lindsey calls on his fellow classical liberals to abandon their “fusionist” political alliance with conservatives and instead seek common ground with the progressive left. After noting some overlap between “libertarian” and “progressive” policy stances, he says that a nascent alliance between these groups should be one that, “at the philosophical level, seeks some kind of reconciliation between Hayek and Rawls”.<sup>160</sup> A colleague of Lindsey and ideological fellow traveller, Will Wilkinson, would dub this nascent philosophical view, “Rawlsekianism,” the aim of which is “to fortify Rawls’ theory of justice with a Hayekian grasp of the coordinating function of prices, and the dynamics of spontaneous order”.<sup>161</sup> As a political movement, for better or for worse, the movement does not appear to have gotten very far. However, several contemporary classical liberal philosophers are working in a “Rawlsekian” framework as it is defined by Wilkinson. Jason Brennan and John Tomasi have given this philosophical framework the more formal academic title of “neo-classical liberalism”.<sup>162</sup> It behoves this thesis to explain why it has not adopted this framework.

---

<sup>160</sup> (Lindsey 2006)

<sup>161</sup> (Wilkinson 2006)

<sup>162</sup> (Brennan and Tomasi 2012)

This chapter thus considers two of the leading attempts to reconcile Rawls and Hayek, those of John Tomasi and Gerald Gaus respectively. Tomasi's approach is based on a sharp distinction between institutional rules and the outcomes – the “order of actions” in Hayek's terminology – which they engender. On Tomasi's reading of Hayek, we cannot directly engineer the order of actions because we lack the requisite knowledge. They are necessarily emergent phenomena. However, we can affect the order of actions *indirectly* via modification of the institutional rules under which agents operate. Tomasi thus reads Hayek's view as being very similar to Rawls, for whom the subject of social justice is the basic structure of institutional rules rather than the outcomes they engender. Tomasi's approach, then, is to socially engineer the basic structure to make maximal use of emergent ordering forces. Thus, collective political agency is still possible under this Rawlsian construal of Hayekian ideas, save that we must, as indeed does Rawls, restrict our focus to the basic rules and not try to directly engineer outcomes. Tomasi terms this approach, “market democracy”.

This chapter criticises Tomasi's view on the grounds that one cannot insulate the design of institutions from Hayek's knowledge problem. It argues against the concept of a societal “basic structure” that can be systemically reformed according to a synoptic conception of justice. Instead, it argues that institutional rule-making and enforcement in impersonal liberal societies is in practice polycentric. There is no Archimedean point from which the basic structure of rules can be reformed. On the view presented here, both the institutional rules and the outcomes which they engender are emergent phenomena. Emergence, we might say, runs all the way down. Gerald Gaus takes a different tack to Tomasi. He, like this chapter, sees both institutional rules and the outcomes they engender as emergent phenomena. However, Gaus tries to claim that institutional rules can emerge spontaneously through human action and not design but nonetheless be seen as a collective political choice. For Gaus, we have consented to rules for whose spontaneous emergence and entrenchment our actions were causally necessary even if we did not *intend* those rules to be dominant. This thesis chapter criticises Gaus' view by arguing that the unintended nature of emergent phenomena cannot be squared with the intentionality necessary for genuine consent.

The argument to follow is divided into six sections. Since the reconciliatory approach stems in part from the limitations of Hayek's critique of social justice, the first section begins with a brief summary of that claim and the debate it inspired. Hayek's critique of social justice did little to question the necessity and possibility of institutional design and focused its ire on those understandings of social justice that aimed directly for patterned distributive outcomes. Section one concurs with Hayek's critics that his critique has little to say to



conceptions of social justice focused on institutional rules and situates Tomasi's fusion of Hayek with Rawls within the context of that debate. However, sections 2-5 try to show that Hayek's knowledge problem can be applied to questions of institutional design in a manner that poses a more fundamental challenge to the Rawlsian framework. To this end, the second section shows how Rawlsian justice presupposes a monocentric source of authority that can act upon and reform institutional rules holistically. The third and fourth sections critique this assumption. The third argues that the state is not the only source of major institutional rules in the social order. Private governance and informal social norms are among the other sources of rule-making and -enforcing authority. The fourth argues that the state itself is better seen a network of partially independent agencies rather than a tightly controlled hierarchy that acts and speaks with one legislative voice. The upshot is that there is no agency that can holistically and systemically reform institutional rules in the manner presupposed by the Rawlsian framework. Section four closes by showing that this critique undermines even John Tomasi's classical liberal construal of Rawlsian justice.

Sections five and six turn to Gerald Gaus' version of the reconciliation strategy. Section five details Gaus' argument that spontaneously emergent institutional rules have been consented to by those whose actions were casually necessary for their emergence. Section six critiques Gaus' argument and also responds to two responses he might make. The nub of our critique is that Gaus conflates Humean acquiescence with more demanding notions of consent and intentional agency. Finally, I do not claim the construal of Hayekian ideas presented here to be exegetically accurate of Hayek's own views. Those who see Hayek and Rawls as fellow travellers can muster ample exegetical evidence that Hayek himself was on their side, though in the latter part of his career he evinced greater scepticism about our ability to design liberal institutions, as we saw in the previous chapter. The argument of this chapter is closer to the spirit of the later Hayek. Still, this is a *Hayekian* argument that departs from some of Hayek's own claims and works.

### **1. The Limitations of Hayek's Critique of Social Justice**

Hayek's critique of social justice in the second volume of *Law, Legislation and Liberty, The Mirage of Social Justice*, is directed principally at one type of such theory, namely, one which requires the distribution of income to track the moral desert of persons. This is peculiar target given that such theories are not the only or even the most prominent accounts of social justice within the academy. I believe Hayek's critique was aimed at more popular

intuitions one finds in conventional political discourse. What he had in mind, I think, was the popular complaint one hears in market societies that it is unfair that members of morally upright professions, teaching or nursing say, earn so much less than professions considered less upright, such as banking or estate agency. Hayek considers these intuitions to be atavistic holdovers from the ethic of the small tribal group. In such proximate settings, where we can reliably predict the impact our actions will have on our fellows and where we have concrete knowledge of their conduct, it is feasible to enforce a conception of justice patterned on desert.<sup>163</sup> In a large-scale social order, however, we lack information as to the merit of most of our peers.<sup>164</sup> Therefore, we lack the epistemic capability to implement patterned conceptions of justice. At bottom for Hayek, such sentiments reflect a misunderstanding of the emergent nature of market distributions and a naïve, anthropomorphic belief that incomes must be consciously decided by a central agency.<sup>165</sup>

This is not to say that market outcomes will be arbitrary. Well-functioning markets will tend to allocate income to those who utilise their abilities in ways that are useful to others.<sup>166</sup> However, the usefulness of one's talents to others in a diverse society will at best weakly track their moral worthiness, if "worthiness" here means the degree of effort one has put in or hardship they have experienced. In David Schmidtz's characterisation of Hayek's view, a free society rewards outputs rather than inputs.<sup>167</sup> In addition, since we lack the capacity to directly determine market outcomes, Hayek claims that it is conceptually incoherent to evaluate them as just or unjust:

Strictly speaking, only human conduct can be called just or unjust. If we apply the terms to a state of affairs, they have meaning only in so far as we hold someone responsible for bringing it about or allowing it to come about. A bare fact, or a state of affairs which nobody can change, may be good or bad, but not just or unjust.<sup>168</sup>

From this premise it follows straightforwardly that the emergent distributions of income we see on the market cannot be just or unjust. If ought-implies-can but we cannot directly engineer market distributions, then social justice so understood is indeed a mirage.

There has been a voluminous literature on Hayek's critique. It is fair to say that his view has generally not been well received. Some of the criticisms, I believe, are misplaced. Many argue

---

<sup>163</sup> (Hayek 1982b, 90)

<sup>164</sup> (Hayek 2006 [1960], 83)

<sup>165</sup> (Hayek 1982b, 62–64)

<sup>166</sup> (Hayek 2006 [1960], 82–85)

<sup>167</sup> (Schmidtz 2016)

<sup>168</sup> (Hayek 1982b, 31)

against Hayek by pointing to cases where *individuals* have *ex-post* duties towards others who have been wronged or are in distress. Under such circumstances the duty-bearer has obligations to assist despite bearing no responsibility for the original suffering or distress. Thus Raymond Plant points out that a bystander acts unjustly if they do not help an elderly person who has fallen into the gutter on a windy day.<sup>169</sup> Eric Mack observes that one can come to possess stolen goods unknowingly but still have duties of rectification to the owner.<sup>170</sup> I do not think these criticisms hit the mark. Hayek is not denying that I can sometimes be held responsible to change state of affairs which I did not intend or whose occurrence I had nothing to do with. Hayek's claim is that I can only be responsible for things that are in my *capacity* to change. These individually-focused case studies clearly show duty-bearers with such a capacity. They do not refer to "state of affairs which nobody can change". By contrast, Hayek's claim is that there is no agent or agency *capable* of directly engineering distributive outcomes on the market to accord with desert, even one who reacts *ex-post*.

A more substantial difficulty for Hayek's view, however, is the aforementioned fact that not all or even the most prominent theories of justice aim to directly bring about a particular distribution. This is certainly not the Rawlsian strategy. Rawls famously says that his theory of justice is an instance of pure procedural justice – "there is a correct or fair procedure such that the outcome is likewise correct or fair, *whatever it is*, provided that the procedure has been properly followed".<sup>171</sup> He goes on to issue the following remark that could forgivably be mistook as one of Hayek's: "[i]f it is asked in the abstract whether one distribution of a given stock of things to definite individuals with known desires and preferences is better than another, then there is simply no answer to this question".<sup>172</sup> The site of Rawlsian justice is the basic structure under which private decision-making takes place. If the basic structure is just so *ipso facto* is any resultant distribution. Indeed Hayek himself claims to be in fundamental agreement with Rawls' approach to justice. In *The Mirage of Social Justice* he writes that he has "no basic quarrel" with Rawls for precisely this reason.<sup>173</sup> Once Hayek makes this concession, however, his critique of social justice, if construed more broadly than the narrow target of principles aiming for a direct distribution, appears to lose its bite. What's more, as we saw in chapter one, Hayek himself concedes that we can make broad-based if

---

<sup>169</sup> (Plant 1994, 171)

<sup>170</sup> (Mack 2018, 98–99); for other variants of this critique see Johnston (1997, 610) and Lister (2013, 412–13)

<sup>171</sup> (Rawls 1999, 75); emphasis added.

<sup>172</sup> (Rawls 1999, 76)

<sup>173</sup> (Hayek 1982b, 100)

not fine-grained predictions about the operation of different institutional rules. This belief is reflected in the many proposals in his own writings to improve the spontaneous workings of liberal society by reforming its institutional rules. If Hayek concedes that we are capable of reforming institutional rules and have at least some ability to predict the effects of those changes, then he appears to offer no categorical critique of social justice as understood in the dominant Rawlsian sense, an observation made by many critics.<sup>174</sup> Indeed, in one of his discussions of Rawls he even concedes that there “exists a genuine problem of justice in connection with the deliberate design of political institutions”. In the same passage he also offers a thought experiment extremely similar to the original position as a device with which we can engage in comparative institutional evaluation.<sup>175</sup> It seems, then, that Hayek is a Rawlsian after all.

Now, if this is accepted, it doesn't follow that Hayek has nothing to say to Rawls. The Hayekian can accept the general Rawlsian framework while tempering some of Rawls' specific proposals on the basis of Hayek's epistemic concerns.<sup>176</sup> It is against this backdrop in the academic debate that, I think, we can understand John Tomasi's insightful proposal to combine a Rawlsian justificatory approach to social institutions with a Hayekian appreciation of the importance of spontaneous ordering in achieving social goals.<sup>177</sup> He terms this proposal “market democracy”. His proposal amounts to a classical liberal construal of Rawls' conception of justice as fairness. The central proposition is that economic liberties should be regarded as basic and consequently deserving of strong constitutional protection along with the civil and expressive liberties normally favoured by left-liberals. Though there are also other policy programmes.<sup>178</sup> Running throughout Tomasi's analysis is a Hayekian scepticism about the effectiveness of securing social goals directly through institutional guarantees.<sup>179</sup> Instead, he argues that Rawlsian aims are best attained by the indirect method of securing basic liberties, including economic ones, and allowing spontaneous ordering forces to

---

<sup>174</sup> Versions of this objection have been advanced, among others, by Miller (2003, 108), Lister (2013, 423) and Lukes (1997, 78).

<sup>175</sup> (Hayek 1982b, 132)

<sup>176</sup> For example, the lack of counterfactual knowledge means that we'll never truly know if the difference principle has been met.

<sup>177</sup> I refer here to his *Free-Market Fairness* (2013)

<sup>178</sup> These include government funding (though not provision) of education and health care (2013, 243-244), constitutionally enforced constraints on government borrowing and fixed monetary expansion rules (2013, 255), free trade and open immigration (2013, 261-262), a basic income (2013, 230) and even possibly affirmative action (2013, 310).

<sup>179</sup> See particularly Tomasi (2013, chap. 7).

operate.<sup>180</sup> Spontaneous order is, in Tomasi's words, a "strategy of social construction" in achieving social justice.<sup>181</sup> Far from being incompatible, Hayekian and Rawlsian liberalisms are complementary on Tomasi's score.

The inadequacy of Hayek's own critique of social justice notwithstanding, his knowledge problem poses a deeper challenge to the associational understanding of political society than implied by Tomasi's proposal. In unearthing this challenge, we can again make note of Hayek's changing attitude towards deliberate institutional design as his career progresses. Compared with his early work, the Hayek of *Law, Legislation and Liberty* and *The Fatal Conceit* is much more likely to describe the rules of just conduct as themselves emergent phenomena. The remainder of this chapter will build on Hayek's later scepticism. To say that social justice is not about the distribution of goods and services to particular persons but the design of institutional rules merely pushes back the epistemic worry. Given the fragmentation of knowledge between numerous and heterogeneous actors in large-scale societies, why should we think that the complex structure of rules and norms under which they interact can be holistically designed to effect social justice? We shall develop and articulate this argument in sections 2-4.

## 2. The Implicit Monocentricity of the Basic Structure

Firstly it is important to see that the Rawlsian idea of a basic structure presumes the existence of an authority with the capacity to reform the major institutional rules of a society systemically. That is the task of this section. I borrow from North, Wallis and Weingast the following broad definition of institutions.

Institutions are the "rules of the game" . . . the patterns of interaction that govern and constrain the relationships of individuals. Institutions include formal rules, written laws, formal social conventions, informal norms of behaviour, and shared beliefs about the world, as well as the means of enforcement.<sup>182</sup>

What is helpful about this definition is that it includes not simply formal laws but also extends to conventions and norms. The importance of this will become apparent in the argument to follow.

---

<sup>180</sup> (Tomasi 2013, chap. 8)

<sup>181</sup> (Tomasi 2013, xvii)

<sup>182</sup> (North, Wallis, and Weingast 2009, 15)

The defining characteristic of Rawls' "basic structure" is that it refers to the major institutions of a society understood as a *unitary system*.<sup>183</sup> As we saw in chapter one, examples of the institutions that comprise such a system include "the legal protection of freedom of thought and liberty of conscience, competitive markets, private property in the means of production, and the monogamous family".<sup>184</sup> This emphasis is not incidental to Rawls' project. It is the manner in which these institutions *combine* with one another to produce a profound, pervasive and crucially *differential* effect on citizen life chances that is important for Rawls. If we are worried about the causes of such inequality, we cannot simply focus on the operation of the education system in isolation from, say, the distribution of healthcare and in turn we cannot think about these without considering, say, the operation of the tax system. Unequal life chances are a product of how the system of rules operates rather than simply the operation of its bare institutional components. Our efforts at institutional reform, by the same token, should be systemic. Piecemeal and disconnected interventions on specific parts of the system would be insufficient. This is the idea Rawls is trying to express when he says that the basic structure is the primary subject of a conception of justice. Indeed, he says that justice as fairness sets up an "[a]rchimedean point for assessing the social system".<sup>185</sup> Tomasi appears to agree with Rawls that a conception of justice applies to "the social order as a whole".<sup>186</sup> This is a task for which he thinks a classical liberal conception of justice is better suited than those preferred by Rawls and other left-liberals.

This Rawlsian aspiration, however, is feasible only if we assume a central pole of authority that is capable not just of reforming all major institutional rules but doing so in a coordinated manner. Each intervention will be but one step in a larger plan to make the system of rules accord with the wished-for conception of justice. To borrow a term from Vincent Ostrom, it presupposes a monocentric source of institutional rule-making and reform.<sup>187</sup> Against this, I wish to argue that rule-making authority in large-scale societies is fundamentally polycentric. There are many sources of institutional rules. Since no one authority has a monopoly on their production, none have the power of wholesale reform over the system of institutional rules holistically understood. Further, Hayek's knowledge problem places stern limits on the ability of these centres of authority to directly coordinate their rule-making activities such

---

<sup>183</sup> It refers to the manner in which major social institutions are "taken together as *one scheme*" (Rawls 1999, 6); or "fit together into one *unified system* of cooperation" (Rawls 1996, 22); or "fit together into *one system* of social cooperation" (Rawls 2001, 10). Emphasis added.

<sup>184</sup> (Rawls 1999, 6)

<sup>185</sup> (Rawls 1999, 230–32)

<sup>186</sup> (Tomasi 2013, xiv)

<sup>187</sup> (Ostrom 1972, 3–5)

that they accord with a normative blueprint. Since the system of rules is the product of the rule-making authority of many actors and sources, it cannot be the subject of any synoptic design.<sup>188</sup>

Some further clarifications are in order at this stage. We should distinguish between the claim that we can normatively *evaluate* a system of institutional rules and the claim that we can *comprehensively reform* them. Hayek appears to deny the former claim in his later work.<sup>189</sup> This chapter does not. We can safely evaluate the system of rules in Sweden as much better than those in North Korea. This is not to deny however, as scholars in the discipline of development have highlighted, that identifying institutional systems that are better than the status quo and successful reform in that direction are two quite different things.<sup>190</sup> For ease of exposition, I will from now on just refer to “institutions” or “institutional rules” rather than adding the “holistic” or “systemic” adjectives. The over-arching focus on the system of rules should make clear that particular agents or agencies are capable of changing particular rules and these decisions often have a predictable impact on particular actors. Barnet Borough Council has recently increased its council tax rates and as a resident of the Borough of Barnet I can personally testify that the decision has had an impact. I am also not denying that the decisions of some actors can have a great deal more impact on the system of rules than others. The decisions made by the House of Commons has much more impact on the system of institutional rules of Great Britain than those made by Barnet Borough Council. Still, the claim is that there is no one locus of agency that can alter the whole system of rules as it likes.<sup>191</sup>

---

<sup>188</sup> I offer then an Ostromian construal of Hayek’s ideas. While Michael Polanyi coined the term, the notion of a “polycentric order” of overlapping but independent rule-making jurisdictions was developed extensively by the Ostroms e.g. Ostrom (1990) and Ostrom (1972). The difference is that the Ostromian project, as I understand it, was to argue for the empirical and normative benefits of institutional alternatives that are more polycentric than prevailing structures. I am instead emphasising that already-prevailing institutional structures are more polycentric than is often assumed.

<sup>189</sup> In the second volume of *Law, Legislation and Liberty* (1982b, 24) Hayek claims that rational evaluation can only be “immanent” to a particular system of rules. That is, our rational evaluation is ever limited to evaluating a proposed new rule or a proposed amendment to an existing rule on grounds of its compatibility with the other pre-existing rules in the system. On this strongly coherentist view, one lacks a rational basis to step back and rationally evaluate the system of rules taken as a whole.

<sup>190</sup> On this see Easterly (2007).

<sup>191</sup> I am grateful to Selina Hofstetter whose forceful objections made me clarify this point.

The monocentric source of authority presumed by Rawlsian and other advocates of comprehensive institutional change is the state. David Miller puts the point well when he writes that theories of social justice presume:

... that there is some agency capable of changing the institutional structure more or less the way our favored theory demands. It is no use setting out principles for reforming the basic structure if in fact we have no means to implement these reforms. The main agency here is obviously the *state*: theories of social justice propose legislative and policy changes that a well-intentioned state is supposed to introduce.<sup>192</sup>

Further, Rawlsians and others presume that the state-in-question is a liberal democratic one. I will thus spend the next two sections arguing that the capacity of the liberal democratic state to comprehensively reform institutional rules is greatly exaggerated. I will claim in section three that the state is not the only maker and enforcer of fundamental institutional rules and in section four that the state itself is not a unitary agency so much as a network of agencies and groups many of whom have autonomous decision-making power.

### 3. Against Legal Centralism

Many political theorists and social scientists subscribe to the doctrine of legal centralism. This is the view that the state is the *only* important creator and enforcer of institutional rules, without which we could manage at best minimal social coordination.<sup>193</sup> There are strong reasons to think that the doctrine of legal centralism is mistaken. For one thing, any associationally complex society will have multiple rule-making and -enforcing authorities. To varying degrees, homeowners associations, universities, churches, professional bodies, trade unions, corporations, shopping centres, night clubs and so on will create rules that would bind members and those resident on their property. Many will also have internal procedures for settling conflicts and changing the content of their rules. Some of these associations, such as the Catholic Church, are very complex associations indeed.<sup>194</sup> The Rawlsian will immediately respond that the basic structure does not directly refer to the internal governance of such associations but to more fundamental rules that form the background

---

<sup>192</sup> (Miller 2003, 6); emphasis added. I am grateful to Billy Christmas for pointing me to this reference. For an excellent review of how the state is either implicitly or explicitly presumed to be the agent of comprehensive institutional change in contemporary political theory, inclusive of Rawls as well as other theorists, and an argument as to why this is an unjustified methodological presumption, see Christmas and Byas (2019).

<sup>193</sup> (Ellickson 1991, 4, 138–47); (Brennan 2018, 238–40)

<sup>194</sup> I draw here on Jacob Levy's *Rationalism, Pluralism and Freedom* (2016a).



against which all associations and persons operate. Examples might be rules demarcating property rights, contract law and the provision of public goods. Surely, the Rawlsian will object, these fundamental rules are created and enforced by states and can be revised as necessary in the pursuit of social justice.<sup>195</sup>

Even with respect to these kinds of rules, however, the legal centralist position is weaker than it initially appears. Take the governance of international trade. Peter Leeson documents that 90% of international commercial contracts contain clauses stipulating that any subsequent dispute should be settled by a private arbitration agency. There are over one hundred such agencies ranging from the London Court of International Arbitration to the Arbitration Institute of the Stockholm Chamber of Commerce. He also observes that there is no formal or universal body of international trade law. The parties have to select the body of law with which any dispute will be settled. These range from the evolved customary law of the *Lex Mercatoria* to any national body of law.<sup>196</sup> Thus, there is neither a single source nor enforcer of the rules governing international trade. Edward Stringham has documented how in both the past and more recent times the basic rules governing market exchange have often been produced by private actors. The rules governing the first stock exchanges in Britain and Holland were formulated and enforced endogenously by the exchanges themselves since state actors would refuse to enforce what they saw as immoral gambling.<sup>197</sup> Today, many consumer e-commerce transactions are facilitated by the privately produced rules and regulations of financial intermediaries such as Paypal, Visa or Ebay.<sup>198</sup> International financial markets also rely a great deal on privately produced rules. For example, the contractual rules for complex over-the-counter financial derivatives are created by an international trade organisation, The International Swaps and Derivatives Association.<sup>199</sup> Stringham is also sceptical of the “shadow-of-the-state” thesis that these rules are enforceable only because the parties can ultimately turn to the state. As with the early stock markets, states sometimes refuse to enforce privately produced rules. Other times, state actors do not know how to enforce rules governing markets of which they have little knowledge. Still at other times the

---

<sup>195</sup> Such is the view developed and defended by Murphy and Nagel (2002) in their *The Myth of Ownership: Taxes and Justice*.

<sup>196</sup> (Leeson 2008)

<sup>197</sup> (Stringham 2015, chaps. 4 & 5)

<sup>198</sup> (Stringham 2015, chapter 7); of course there are markets which rely on state-enforced law. Those markets reliant on stringent intellectual property rights might be an example. The point here is that this is not as general a truth as often assumed.

<sup>199</sup> (Stringham 2015, 179–83)

volumes of transactions are so high but the unit value of each is so low that it is impractical to turn to the state for every instance of wrong-doing.<sup>200</sup>

Robert Ellickson has documented case studies of market actors governing themselves through informally generated social norms. These actors range from international whalers to Californian ranchers.<sup>201</sup> He suggests that social and commercial life might be governed by informal norms to a much greater extent than one might assume from a legal centrist perspective. He cites empirical research showing that ordinary people “know little of the private substantive law applicable to decisions in everyday life”. For example, 96 per cent of Californian therapists did not know of an important California Supreme Court ruling that dealt with therapist tort duties in instances where their patients had uttered threats against third parties.<sup>202</sup> If market agents are often unaware of the legal rulings that are supposed to govern their exchanges, this suggests that their impact might be more limited than is often assumed. As for public goods, Elinor Ostrom has famously documented many cases of communities around the world establishing and enforcing their own rules with respect to the provision of local public goods.<sup>203</sup> Even with regard to the most basic of state functions, we find that states do not always have a monopoly on the creation and enforcement of institutional rules. David Skarbek has documented how the United States has effectively outsourced the internal governance of most of its prison system to prison gangs, many of whom have their own explicit constitutions to which new members have to agree upon joining.<sup>204</sup> None of this is meant to deny that state law-making has a very substantial impact on the composition of institutional rules in any social order. Rather, to borrow a phrase from Jason Brennan, the “marginal value of public governance” is much less than one might assume.<sup>205</sup>

It might be objected that these examples merely show instances where the state has unnecessarily ceded power to other actors. All they demonstrate is the need for stronger state institutions that can “rationalise” the social space by bringing these plural and disparate source of authority under its control.<sup>206</sup> However, there is an epistemic reason to think that state actors will have to cede significant amounts of *de-facto* decision-making power to other

---

<sup>200</sup> (Stringham 2015, chap. 1)

<sup>201</sup> (Ellickson 1991)

<sup>202</sup> (Ellickson 1991, 145)

<sup>203</sup> (Ostrom 1990)

<sup>204</sup> (Skarbek 2011)

<sup>205</sup> (Brennan 2018)

<sup>206</sup> The notion of the rationalising state that incorporates intermediate sources of authority is explored by Jacob Levy (2016a) in his *Rationalism, Pluralism and Freedom*.

actors. The production of institutional rules for complex societies faces a knowledge problem analogous to that facing economic coordination.<sup>207</sup> Different domains of social life require different rules depending on local circumstances and the state is often poorly placed to assess these compared with actors on the ground. Private arbitration, to take an example, is popular in complex commercial contracts because the parties typically require agencies who have specialist knowledge of their business sector. Universities decide upon and enforce their own scholarly standards because they, rather than state officials, have the best understanding of them.<sup>208</sup> Further indication of this epistemic deficit comes from the fact that when states do appropriate the authority of non-state bodies, they often replicate the institutional rules already built up by associations and practitioners on the ground. When the US Federal Government implemented the Securities and Exchange Act of 1933 and 1934, for example, it “simply mandated many of the requirements that the New York Stock Exchange had already adopted”.<sup>209</sup> David Green and David Beito demonstrate that many of the functions performed by the contemporary welfare state were already pioneered by 19<sup>th</sup> century worker-run mutual aid societies.<sup>210</sup>

It is also worth taking further note of informal social norms as a particularly important alternative authority to the state. The topic of social norms has gained prominence among political theorists and social scientists in recent years. Christina Bicchieri’s high profile account defines social norms as default rules of behaviour to which agents conform because they believe that a sufficiently large subset of their peers follow the norm and that a sufficiently large subset of their peers normatively expect them to follow it.<sup>211</sup> Such social norms, then, are enforced in the first instance not by the threat of legal sanction by the reciprocal expectations of compliance by community members. There is no doubt that states sometimes have the power to change social norms by legislative action. The rapid change in social norms with respect to drink-driving following state campaigns might be a good example of this. More often than not, however, social norms constrain the legislative ability of states such that laws in tension with social norms are not widely followed. Thus, Bicchieri observes that a range of programmes in the developing world, from those designed to prevent child marriage to educating girls to inducing communities to build and use latrines

---

<sup>207</sup> Here we are combining Hayek’s insights with those of Ostrom (1990).

<sup>208</sup> To take a more concrete example, Stringham documents how the private firm PayPal had to develop its own risk-management tool to deal with fraud simply because officials at the FBI had very little knowledge of the internet or their business model. (Stringham 2015, chap. 7)

<sup>209</sup> (Stringham 2015, 81)

<sup>210</sup> (Green 1993); (Beito 2000)

<sup>211</sup> (Bicchieri 2006, 5)

have failed because they were not in keeping with local social norms.<sup>212</sup> The failure of various Western states' attempts to wage "wars" on drugs might be interpreted in a similar light. The prohibition against drugs cuts against the social norms of large swathes of the population.

The argument thus far is consistent with Adam Smith's famous critique of the "man of system". He thinks "he can arrange the different members of a great society with as much ease as the hand arranges the different pieces upon a chess-board" but does not consider that on "the great chess-board of human society, every single piece has a principle of motion of its own, altogether different from that which the legislature might chuse to impress upon it".<sup>213</sup> The man of system overlooks that he is but one of many rule-making and -enforcing authorities in the social order. These other agencies have a "principle of motion" of their own. This alone, I think, is enough to undermine the Rawlsian thought that we can comprehensively reform the basic structure from some Archimedean point. However, we can go further than this. The idea of a singular *man* of system is itself an abstraction. The state is not a unitary actor but closer to being a polycentric network of partially independent agencies and actors.

#### 4. The State as a Polycentric Network

Consider three ways we might model the state:

1. *Leviathan* – a singular and all-powerful actor.
2. *Corporate organisation* – a hierarchically-structured organisation led by a centralised decision-making body i.e. a legislative chamber.
3. *Network* – an interrelated and decentred network of agencies, each with significant discretionary power and autonomy.

The Leviathan model, dating back to Hobbes, is a popular way of modelling the state, manifestations of which include the "revenue-maximizing monarch, a stationary bandit, or a single-actor 'representative agent'".<sup>214</sup> We might conceive of such an agent as an all-

---

<sup>212</sup> (Bicchieri 2006, 48)

<sup>213</sup> Cited in Levy (2016a, 66–67). I here follow Levy's reading that the "pieces" with their own "principle of motion" are not the bakers, brewers and butchers from Smith's other famous metaphor but distinct rule-making authorities.

<sup>214</sup> (North, Wallis, and Weingast 2009, 17)

powerful actor who can intervene exogenously upon the system of rules. It is probably this understanding of the state that is most conducive to the Rawlsian aspiration of wholesale institutional reform. Whatever its theoretical utility in certain contexts, it is neither sociologically nor empirically realistic and I think we can safely discard it here. I suspect that most political theorists have the organisational model in mind. This is certainly more empirically realistic. We picture an organisation at the apex of society with legislative authority to reform existing institutions. Note that the state as organisation is already constrained in its ability to effect reform compared with the state as leviathan. Such a state is no longer an all-powerful entity exogenous to the institutional rules of the game but a bureaucratic organisation whose operations are, in part, constrained by those rules. Its operation might be constrained by a separation of powers for instance. To the extent that it is so constrained, its ability to reform the institutions holistically will be so constrained. I however wish to go further and claim that the state itself is closer to a network of actors than a unitary organisation controlled by a legislative centre.<sup>215</sup>

The term “state” is a short hand for a network of different agencies. These agencies are *functionally* differentiated, consisting of executive, legislative, judicial, administrative, defence and welfare bodies. Internally each of these bodies will be further differentiated into sub-functions. For example, the judicial system in the United Kingdom consists of the Supreme Court, criminal courts, civil and family courts and so on. State agencies are also *hierarchically* differentiated. There are local, central and indeed international government bodies. For example, citizens of the United States can be subject to the laws of the United States Congress, state legislatures, city governments, municipal governments and county governments, not to mention native-American tribal governments. The welfare or regulatory arms of contemporary states, particularly, consist of a vast number of agencies and organisations. Under the rubric of the “British State” one can find the National Health Service, the British Broadcasting Corporation, The British Military, the Department for Education, The British Police force, the Secret Intelligence Service or “MI6”, The Security

---

<sup>215</sup> A couple caveats are in order at this juncture. First, the analysis to follow is restricted to liberal democratic states chiefly because that is the sort of institution that Rawlsians and others think should implement social justice. Though much of the analysis below is applicable also to non-democratic and/or non-liberal states whose activities are large in scale and complex in scope. Second, the argument to follow is applicable to state institutions as we find them in typical nation states such as the United Kingdom or the United States whose populations run into the millions and in which the scale of state activity is commensurately large. It may not apply to smaller states such as Monaco or the Vatican. At the same time, Rawls and other advocates of social justice also have these larger states in mind and do not restrict their claims to smaller outlier states.

Service or “MI5” and The Bank of England just to name a few.<sup>216</sup> There is a reasonable case also that we should include among the agencies of the state some nominally private entities. The distinction between “public” and “private” is not always clear-cut. For example, when states outsource or privatise a function but continue to regulate the entities that take over those functions, the status of those entities becomes unclear. It is not obvious whether, say, firms in the United Kingdom’s highly regulated private rail network are “private” or “public” organisations. However, the analysis to follow will focus on “formal” state agencies, thus it will even downplay the complexity of the network of state institutions.

Several considerations can be offered in support of the contention that these various entities are closer to a network of agencies, each with a “principle of motion of its own”, than a unified hierarchy controlled by a legislative centre. The first is democracy. At any one point in time, different branches of the state will be composed of the representatives of different parties. Over time, the composition of different branches will change from election-to-election. Whether viewed at a particular moment-in-time or diachronically over time, the legislative output of states will be a function of the representatives of different political parties with often incommensurable if not directly opposed aims. Nor can we say that the outputs of legislative chambers straightforwardly reflect the intentions of any actors. As Jeremy Waldron has observed, the divergent and incommensurable views of individual legislators means that they often can’t directly agree on the content of legislation. Instead, they fall back on procedural rules that can yield collective legislative outputs from the variegated inputs of individual legislators. The upshot, however, is that “under the conditions of modern legislation, it is often implausible to describe legislative acts as intentional acts”. Instead we should see legislation as an “artificial resultant of the acts of individual legislators, structured and related to one another through certain procedures, decision-functions and perhaps machines”.<sup>217</sup> Waldron’s radical thesis is that much contemporary state legislation is the result of the human actions of many legislators but not any overall design.<sup>218</sup>

---

<sup>216</sup> According to the UK government website, the British state consists of 25 Ministerial Departments, 20 non-ministerial departments, 402 Agencies and other public bodies, 82 High Profile Groups, 12 Public corporations and 3 devolved Administrations. That is a total of over 500 distinct bodies and department (UK Government Website, 2019).

<sup>217</sup> (Waldron 2004, 128)

<sup>218</sup> It is worth noting that the dynamic of competition for the reins of power is hardly limited to democratic states. Even nominal one-party states consist of coalitions with differing aims and interests who frequently compete with one another. They are not unitary entities. The Chinese Communist Party, for instance, is often understood to consist of two distinct “elite” and “populist” coalitions. Purges, whereby one part of the elite coalition removes another, is not an infrequent occurrence in

Another relevant consideration is the manner in which state institutions are deliberately designed to have firewalls between one another and prevent concentrations of power. The most obvious example in this regard is the separation of the legislative and judicial functions of the state. However, state institutions are deliberately fractured along other dimensions. Actors within state institutions will often deliberately cede authority to other actors in order to make credible commitments to third parties. Over the last few decades the governing parties of many liberal democratic states ceded decision-making capacity with respect to interest rates to independent central banks. In tying their own hands from using interest rate policy for short-term electoral purposes, they could make credible commitments to third-party economic actors in financial markets and elsewhere. Political economists Douglass North and Barry Weingast explain the formation of liberal democratic institutions through just such a mechanism.<sup>219</sup> The British Crown was comparatively much better placed to raise debt finance following the introduction of parliamentary constraints on its sovereignty in the aftermath of the Glorious Revolution. These constraints reassured the King's creditors that they had credible means to hold the King accountable to his financial commitments. In turn, they were more willing to lend to the Crown. When the actors in the "centre" of what we might call the "state-network," be they a monarch or an elected government, make such commitments, they will be credible only if hands genuinely are tied and decision-making powers genuinely devolved to other agencies, as indeed seems to have been the case in the Glorious Revolution.

The third consideration is epistemic. The sheer scale and breadth of contemporary state activity is quite remarkable. A subset of the activities undertaken by contemporary states include war-making, combatting terrorism, regulating drug use, promoting cultural integration, devising pedagogical techniques to teach children, administering tax schemes, combatting climate change, regulating nuclear waste, administering welfare, and so on. Many of these activities require specialist knowledge that cannot be centralised but will be held by practitioners within the relevant agency. For example, the decisions as to which health treatments are available on the United Kingdom's National Health Service are made not by MPs but by a specialist clinical body, the National Institute for Health and Clinical Expertise.<sup>220</sup> The vast majority of members of the UK Parliament are not trained medical professionals and lack the knowledge to make these assessments. Crucially, the

---

such states. For an argument that anticorruption drives are used in China as weapons in intra-elite power struggles, see Zhu and Zhang (2017).

<sup>219</sup> (North and Weingast 1989)

<sup>220</sup> (Rawlins and Culyer 2004)

fragmentation of knowledge here is not only of theoretical knowledge but also the more practical sort highlighted by Hayek. What Mill observed of the nineteenth century British civil service is no less true today:

Every branch of public administration is a skilled business, which has its own peculiar principles and traditional rules, many of them not even known in any effectual way, except to those who have at some time had a hand in carrying on the business, and none of them likely to be duly appreciated by persons not practically acquainted with the department.<sup>221</sup>

The less the legislative centre of the state-network understands the activities and practices of agencies at the “periphery”, the less it can directly control their activities.<sup>222</sup>

The fourth consideration is based upon an insight of Carl Schmitt. He argues that the implementation of a rule has an unavoidably “decisionistic and personalistic” element to it.<sup>223</sup> His point is not merely the familiar one that law is indeterminate and requires interpretation. Schmitt says *who* interprets the rule matters. That one interpretation rather than another will be selected is a function of the personality, characteristics, judgements, circumstances and so forth of the particular official. If we interpret Schmitt through a Hayekian lens, we can say that the substantive interpretation of a law depends on the local knowledge and values of the *in-situ* practitioner. Even when the state-network appears to operate in the most hierarchical fashion, the centre might be in less control than it appears because the “gaps” in the formal instructions it issues are filled by other state agents. This was precisely what Michael Lipsky found in his famous study of “street-level bureaucrats”. Central state policies would list several criteria or values to be upheld without specifying how they were to be traded-off against one another. These and other such decisions were left to the street-level bureaucrats who tangibly interacted with citizens on the ground. Lipsky concludes that such bureaucrats had significant discretion in deciding how general and often vague policy commitments emanating from the centre were concretely implemented.<sup>224</sup> If the implementation of a policy depends to a non-negligible degree on the discretion of agents on the periphery of the state-network, this is another constraint on the ability of the legislative centre to direct the network.

---

<sup>221</sup> (Mill 1861, 58)

<sup>222</sup> The premise of the British political satire *Yes Minister* was precisely this information asymmetry. The show depicts an elected minister struggling to get his way because he is ignorant of the manner in which departments work in comparison to the civil servants who have spent their whole careers working within them.

<sup>223</sup> (Schmitt 2005, 48); Michael Oakeshott (1991) makes a similar point in his famous essay “Rationalism in Politics”.

<sup>224</sup> (Lipsky 2010)



The fifth consideration is based upon a critique of public policy studies provided by the philosopher Mark Bevir.<sup>225</sup> He observes that the dominant theoretical frameworks used in that discipline, rational choice inspired “neoliberalism” and “institutionalism,” do not adequately account for how the contingent and variable *beliefs* of state actors influences their behaviour. These frameworks assume that such actors’ behaviour is a predictable function of objective and ascertainable social facts. These are narrowly construed incentive structures for the rational choice theorist and fixed institutional procedures for the institutionalist. However, what both these approaches overlook is the manner in which the actors-in-question *subjectively* interpret their situation according to their own lights. An actor might respond differently to a given situation than predicted by the theory or they might even categorise the situation differently than the theory. The upshot is that public policy is not a predictive science. The response of actors in the periphery of the state network to the centre’s directives will depend on the variegated and contingent *interpretations* they make of those directives and their own particular circumstances. This implies that policies will rarely achieve their intended consequences and often generate unintended ones. The outcomes of such policies might be better conceived of as what Charles Lindblom once termed “mutual adjustments” between different agencies. He writes:

In the United States, for example, no part of government attempts a *comprehensive overview* of policy on income distribution. A policy nevertheless *evolves*, and one responding to a wide variety of interests. A process of mutual adjustment among farm groups, labor unions, municipalities and school boards, tax authorities, and government agencies with responsibilities in the fields of housing, health, highways, national parks, fire, and police accomplishes a distribution of income in which particular income problems neglected at one point in the decision processes become central at another point.<sup>226</sup>

Note that for Lindblom “mutual adjustment” does *not* mean that these government or quasi-government agencies consciously decide through an explicit voting or deliberative procedure how income will be distributed amongst themselves. Instead, mutual adjustment takes place between agencies acting upon their own localised information and interpretation of their interests. Hence there is no “comprehensive overview” of the income distribution which instead “evolves” through their iterated interactions. To formulate these insights in the language of this thesis, the income distribution in Lindblom’s example is the emergent outcome of many agencies lobbying for their own purposes and adjusting to one another’s efforts in an unpredictable and complex manner.

---

<sup>225</sup> The following draws upon Bevir (2013)

<sup>226</sup> Lindblom (1959, 85); emphasis added.

The sixth consideration, related to the last, is the dynamic of interest group politics. A significant amount of legislation in contemporary states arises because particular groups have an interest to campaign or lobby for particular laws. In Mancur Olson's well known account, the benefits of such legislation are concentrated among members of the group while the costs are dispersed across society-at-large.<sup>227</sup> Farming subsidies exist, for instance, primarily because farmers are incentivised to lobby for them. To the extent that interest group politics are prevalent in the political process, state legislation will reflect the discrete priorities of different interest groups rather than the coordinated plan of a sovereign legislative centre. The seven consideration is the sheer number of laws and regulations produced by contemporary states. If the state is closer to a polycentric network than a hierarchically structured organisation, we should expect a proliferation of laws on disparate matters that would not exhibit any coherence. This is indeed the case. Michael Huemer observes that the United States Code Of Federal Regulations, which *excludes* statutes passed by Congress as well as state and local laws, ran to 152,456 pages in 2010, covering a plethora of topics ranging from the spacing of spark plug gaps to others describing matters as specific as "the signing of documents related to excise taxes on structured settlement factoring transactions".<sup>228</sup>

We have built up our argument to the following picture. Any social order will consist of a plurality of producers of institutional rules, both within and outside what I have called the state-network. There is no agent or agency who can adopt an Archimedean perspective and intervene holistically on the system of rules as a whole. Such a system is the product of the institutional rule-making and enforcement of a plurality of authorities rather than any overall design. To return to our core contention from chapter one, ethical disagreement or compliance problems are not here the primary difficulties. We could posit that all these disparate actors universally subscribed to, say, Rawls' conception of justice as fairness or Tomasi's conception of justice as free market fairness. The fragmentation of knowledge under large-scale settings is in play here just as it is in the economic context. These authorities produce rules based at least partly on their own proximate knowledge unavailable to the rest. Even positing abstract consensus on a conception of justice, the

---

<sup>227</sup> (Olson 2003 [1971])

<sup>228</sup> (Huemer 2013, 280–81)

problem of coordinating and rendering compatible their rule-making initiatives still remains.<sup>229</sup>

Let me emphasise again that the argument is about our ability to design institutional structures *systemically*. Thus I do not deny that particular political actors are often in a position to reform specific rules. We can even appropriately invoke considerations of social justice in evaluating their decisions. There are actors within the state-network of the United States who might be able to reduce incarceration rates or within that of the United Kingdom who might be able to liberalise onerous restrictions on the supply of new housing. I think one can safely say that both ought to be done and even invoke reasons to do with social justice in both cases. Reforms such as these, however, are piecemeal forms of social engineering in the Popperian sense. They are a far cry from holistic reform of the major institutional rules conceived of as a unitary system along the lines of Rawls' basic structure. I concede that a partial exception to the argument I have offered are certain historical moments when a handful of persons are able to formulate constitutions that subsequently bind the rest of society. The US case most obviously comes to mind. However, such "great constitutional moments" are rare. They might be the exception that proves the more general rule that I am defending here. Even then the ex-post operation and evolution of these institutions will depart from the intentions of the original designers.<sup>230</sup>

At this point, John Tomasi might offer the following objection. Perhaps the arguments mustered in this and the prior section speak against Rawlsian or socialist conceptions of justice that tend to overlook Hayekian concerns. However, Tomasi might argue, it does not speak against his more epistemically modest conception of justice. Free market fairness tends to eschew direct interventionism after all in favour of utilising spontaneous ordering forces as much as possible to further the end of social justice. The hugely ambitious forms of social engineering necessary to realise, say, fair equality of opportunity or the difference principle would not be entailed by Tomasi's construal of Rawlsian justice. In response, it should be conceded that a classical liberal blueprint for the basic structure is probably less

---

<sup>229</sup> The second principle of justice as fairness requires that socio-economic inequalities be arranged so that, first, offices and positions are open to all under conditions of fair equality of opportunity and, second, that their arrangement is of the greatest benefit to the least advantaged. At the least, this seems to require the coordination of multiple layers of government with respect to policy areas as diverse as health, housing, taxation, pensions, monetary policy, trade, the family and education. Even if we posit that all these government officials fully subscribe to justice as fairness, the problem remains as to how they will practically coordinate their many disparate policy initiatives to arrive at the desired outcome given that the fragmentation of knowledge is in play.

<sup>230</sup> Hence the perennial complaint of classical liberals in the US that the scope and scale of the contemporary federal government far exceeds anything envisaged by the founders.

susceptible to the problems we have identified, but it is far from being entirely free of them either. For one thing, as we already observed above, Tomasi's conception of justice does entail the implementation of a policy programme including the funding of education and healthcare, the implementation of monetary rules, the provision of a basic income and even affirmative action. To the extent that these policies add up to a programme of systemic institutional reform then they are susceptible to the concerns we have just raised.

For another, the core of Tomasi's proposal, that state institutions should secure *basic economic liberties*, is definitely susceptible to the concerns raised above. According to Tomasi, economic liberties are liberties of working and owning. Working liberties allow one to use one's time and body in ways that one sees fit. Ownership liberties refer to the liberty to buy-and-sell goods and services as well as to hold and use property.<sup>231</sup> Free market fairness holds that economic liberties should be among the constitutionally protected basic rights of liberal citizens.<sup>232</sup> The basic structure of a Tomasian society, then, is arranged such that all citizens have access to a robust and one would assume mutually compossible list of such liberties among other personal and civil ones. However, we have already seen that some of the most important markets, such as those for international trade or some financial markets, are governed by non-state sources. The economic liberties of actors in these markets are not determined by the state.

Furthermore, it is doubtful whether entrenching a "basic" list of economic liberties for all economic domains is a coherent enterprise in a complex and diverse society. Under such a setting my economic liberties would exhibit at least some variation from context-to-context. Consider the different configurations of economic liberties I have when I buy a complex derivative contract, obtain a mortgage for a house, purchase some trademarked software, farm on land I hold in common with fellow farmers, purchase non-voting ordinary shares in a company, work as a member of an employer-led cooperative, claim the entitlements that I am owed as a member of a mutual society, own real estate concurrently with at least one other person and purchase an asset-backed security whose revenues are derived from the present and future musical sales of leading pop star.<sup>233</sup> It seems very unlikely that we can abstract away from these domains a "unitary list" of economic liberties. Perhaps that is the problem as far as Tomasi is concerned. We can't have a "checkerboard" solution to property

---

<sup>231</sup> (Tomasi 2012, 53–54)

<sup>232</sup> (Tomasi 2013, 121)

<sup>233</sup> The last of these examples is based on a real bond issued by the musician David Bowie.

rights.<sup>234</sup> However, this response would overlook that it is often desirable that the precise configuration of economic liberties or property rights vary from domain-to-domain. What makes for efficient property rules in an Ostromian commons might well not do so in an auction house. In insisting on a unitary scheme of basic liberties, then, Tomasi's conception of justice retains Rawls' overly centralised understanding of institutional rule-making.<sup>235</sup>

## 5. Gerald Gaus' Liberalism

Let us take stock. We began by noting the key limitation of Hayek's critique of social justice. It did not apply to theories of justice, like that of Rawls, whose subject is institutional rules rather than the outcomes they engender. One potential response is offered by Tomasi. Concede the point and try to incorporate Hayekian insights into a Rawlsian justificatory framework. We took another strategy in arguing that Hayek's knowledge problem applies also to the design of institutional rules. Just as there is no Archimedean point from which we can directly engineer distributive outcomes, there is no such point from which we can holistically reform systems of institutional rules. On our construal of Hayekian social theory, both the system of institutional rules and the interactions to which they give rise are emergent phenomena. It behoves us at this juncture, however, to engage with Gerald Gaus' liberal theory.

Gaus, like Tomasi, presents a form of liberal political philosophy that tries to reconcile Hayekian and Rawlsian insights. However, unlike Tomasi, he does not see spontaneous orders only as a tool of social construction that we utilise when consciously designing the basic structure. Gaus thinks that the basic institutional rules of liberal society are themselves the product of an emergent and evolutionary process. On his view, institutional rules are not designed but arise spontaneously through the iterative interactions of countless liberal agents. In this respect, Gaus takes the same approach as this thesis. However, in contrast to this thesis, he also thinks that spontaneously emergent institutional rules can nonetheless be viewed as the product of the shared and active agency of liberal citizens. In other words,

---

<sup>234</sup>This is a reference, of course, to Ronald Dworkin's famously dismissive attitude to competing legal jurisdictions.

<sup>235</sup> I note in this regard the school of thought that sees market institutions arising in Europe spontaneously as a result of competition between many rule-setting authorities. For instance, Barry Weingast (1995) thinks the industrial revolution occurred in Britain as an accidental by-product of its "market preserving federal structure". According to this school of thought, there was no conscious plan to arrange the basic structure to accord with a unitary list of economic liberties. C.f. Pennington (2011, 196–97); (Fink 2012).

that institutional rules are emergent phenomena does not preclude their being understood as chosen and even consented to by the liberal citizens from whose actions they emerge. Rawlsian justificatory ambitions can be combined with a thoroughgoing Hayekian account of the origin of social institutions. Gaus' liberalism therefore provides a distinct and strong challenge to our attempt to drive a wedge between Hayekian and Rawlsian liberalisms. I will argue over this and the next section that Gaus' attempt, while creative, fails. Either the institutional rules have been actively chosen or they are emergent. I will contend they cannot be both and that Gaus' argument plays on an ambiguity between Humean acquiescence and more substantive notions of consent and choice. I draw primarily on his *Order of Public Reason (OPR)*.

Gaus' locates his inquiry on the same pasture as Rawls and Rousseau. His book is offered as an answer to the following question: "How can free and equal moral persons claim authority to prescribe to other free and equal moral persons?" His answer, in short, is that "a publicly justified morality – one that the reason of each endorses – allows each to remain free while subject to moral authority."<sup>236</sup> He is deeply troubled by the potentially authoritarian nature of interpersonal morality. Such morality gives one licence to issue imperative demands to others – "You *must* do this". In issuing a demand, the demander claims authority over another and this is troubling if we assume persons are free and equal.<sup>237</sup> Imagine, however, a community of agents living under moral rules all endorse. Under such circumstances, making a moral demand of another is not inconsistent with respecting them as free and equal. In so doing I am simply asking them to live up to commitments that *ex hypothesis* they themselves endorse.<sup>238</sup>

What is particularly interesting about Gaus' proposal is that his philosophical focus is not simply the judicial laws of the state but social norms or what he calls *social morality* – "the shared rules, shared expectations and requirements, and shared understanding of appropriate demands and sanctions that structure our life together".<sup>239</sup> These are demands that we make upon one another in the context of everyday life. They are not necessarily enforced by the state. Gaus observes that the focus of Mill's *On Liberty* was social morality in this sense i.e. those rules occupying an intermediate position between those coercively enforced by the state and those decided upon by the individual herself on the basis of

---

<sup>236</sup> (Gaus 2012a, 5)

<sup>237</sup> (Gaus 2012a, 22)

<sup>238</sup> (Gaus 2012a, 29)

<sup>239</sup> (Gaus 2015b, 143)

conscience.<sup>240</sup> Legitimate social morality lays the foundation for legitimate state authority on Gaus' score. Coercively imposed institutional rules are legitimate when they accord with or build upon already established and publicly justified social-moral norms. So he is concerned in the first instance with how legitimate rules of social morality arise. The core inquiry of *OPR* is how members of the public might converge upon social-moral rules that all of them can endorse.<sup>241</sup>

The difficulty is that no consensus is forthcoming because liberal citizens have reasonable disagreements in the Rawlsian sense – Gaus uses the term “evaluative pluralism” – about the most appropriate rules. Still, Gaus thinks there are ways in which we can prune this disagreement. With respect to a particular issue, each person can rank rules such that they compile that person's eligible set of rules. This consists of all the rules that person ranks above having no rule with respect to that issue. This allows us to identify a *socially eligible set* of rules that excludes any rule that every member of the public ranks below having no rule on the issue and any rule that is dominated by another rule in the personal eligible set of every member of the public.<sup>242</sup> Thus, if A is ranked higher than B in the personal eligible sets of all, we can drop B from the socially eligible set and if C is ranked lower than no rule in the personal eligible sets of all, we can drop C.<sup>243</sup>

Drawing on Benjamin Constant, Gaus identifies two other further such pruning strategies. Firstly, all members of the public have reasons to affirm those social rules that promote basic rights of agency. Following Stanley Benn, Gaus thinks we moderns, our deep cognitive diversity notwithstanding, cannot but help see ourselves as self-directed agents.<sup>244</sup> *Qua* such agents we can be understood to have reasons to affirm social rules that are reasonable interpretations of certain rights pre-requisite to such agency. These include a defeasible presumption of liberty with respect to the pursuit of our goals, the right not to be coerced or deceived, the right to freedom of conscience, rights against harm and rights to assistance.<sup>245</sup> Secondly, members of the public have reasons to affirm property rights or, to use Gaus' own parlance, jurisdictional rights. Such rights devolve morality on certain matters to individuals and communities. The Catholic Church and the British Humanist Organisation

---

<sup>240</sup> (Gaus 2012a, 3)

<sup>241</sup> Members of the public has a technical connotation in Gaus' theory. They are moderately idealised versions of their real-life counterparts. (Gaus 2012a, 26–27; 267)

<sup>242</sup> The latter is a Pareto rule.

<sup>243</sup> (Gaus 2012a, 321–23)

<sup>244</sup> (Gaus 2012a, 337–40)

<sup>245</sup> (Gaus 2012a, 341–59)

can enforce their own social rules within the jurisdiction of their own property. Jurisdictional rights prune the eligible set by reducing the need to agree on a common rule for many matters. Each member of the public, claims Gaus, will have strong reasons to affirm rules constituting reasonable interpretations of jurisdictional rights so as to uphold their personal evaluative standards in a diverse world.<sup>246</sup> These two further pruning strategies reduce the size of the socially eligible set but do not leave us with a singleton. For abstract rights of agency and jurisdictional rights can be interpreted in many ways. These strategies only exclude those social rules that are not plausible interpretations of these general commitments, still leaving us with a numerous socially eligible set of candidate rules. At this point philosophical reason seems to run out.

Gaus turns to a Hayekian solution. He thinks we can converge on a set of action-guiding rules to govern our social life through bottom-up social interactions. Gaus' argument here is technical and quite abstract. He uses two different game-theoretic models to advance it. I focus here on the model he terms the "Kantian coordination game".<sup>247</sup> The basic idea is to imagine a very large number of agents engaged in repeated interactions. During each interaction agents have the choice of acting on a social rule from the socially eligible set. Agents successfully coordinate if they both act on the same rule. Gaus assumes that agents always prefer successful coordination to their acting on different rules, even if successful coordination requires them to act on a rule that is not their first-best. When the players begin interacting, the initial strategy of each player is to act on the rule that they consider first-best or optimal. However, this means that some players will fail to engage in successful coordination and so will change their strategy to rules they consider less than optimal in subsequent rounds in order to increase their chances of successfully coordinating with others. Over time, Gaus shows formally that the agents will converge on one social rule, which will not be the optimal rule for everyone or even possibly anyone.

The principal mechanism is a network affect. As one rule becomes popular, ever more individuals subscribe to it to increase their chances of successful coordination creating a self-generating tendency towards convergence on that rule. Which rule is selected, however, cannot be predicted *ex ante*. It is the emergent and unintended consequence of individuals trying to act on their most preferred rule given the likewise preferences of all the other players. Despite our converging on the equilibrium rule in "a contingent and accidental way",

---

<sup>246</sup> (Gaus 2012a, 372–86)

<sup>247</sup> (Gaus 2012a, 393–400)



Gaus believes we can still see this as a self-legislated morality in the Kantian/Rousseauian sense.<sup>248</sup> How could this be so? He writes:

With a rule in equilibrium, she [i.e. an agent] best satisfies her standards *given what others are doing*... Her freedom is not the freedom of an asocial agent who is free simply when she does what she thinks best regardless of what others do, but the freedom of a social moral agent who considers what her evaluative standards deem is the best thing to do *given* what others justifiably do on the basis of their own standards. We can think of this as a social Realm of Ends populated by free and equal persons with diverse evaluative standards. When a person takes this perspective, she sees that she is free when she acts on a moral rule that is in equilibrium.<sup>249</sup>

The obvious objection to characterising the equilibrium rule as the product of a “social realm of ends” is that there is no genuine collective choice in these interactions. Gaus acknowledges this – “the group, *as a group*, does not choose its moral rules”.<sup>250</sup> Nonetheless we have chosen the rule through our “interdependent” individual choices:

As interdependent individuals, we choose our moral rules — at least, if we are lucky enough to actually possess bona fide moral rules. In the end, this is a real choice of actual people; not a hypothetical choice, or one dictated by impartial reason.<sup>251</sup>

Since it is a real choice, it binds. Once a rule is converged upon, we can say that “a free society *consents* to the moral rule to which we are all bound”.<sup>252</sup> Gaus’ claim, then, is that a group of persons have chosen to be governed and bound by a rule if they have converged upon it through repeated interactions. This is so even if, by hypothesis, the process of converging on this rule was “contingent and accidental” and the agents involved did not and indeed, due to their disagreements, could not consciously decide upon the rule before their interactions. According to Gaus, then, rules that emerge through human action but not design can nonetheless be described as chosen and even consented to by the actors-in-question.

To verify the accuracy of this reading, I compare it to renderings of the same argument that Gaus has offered in publications separate to *The Order of Public Reason*. In one paper he responds to John Simmons’ famous critique of hypothetical accounts of consent. Simmons has long argued that the generic qualities of a state are insufficient to create an obligation to obey. Actual consent by the governed is also necessary.<sup>253</sup> In response, Gaus reiterates the

---

<sup>248</sup> (Gaus 2012a, 400)

<sup>249</sup> (Gaus 2012a, 402); original emphasis.

<sup>250</sup> (Gaus 2012a, 402); original emphasis.

<sup>251</sup> (Gaus 2012a, 402)

<sup>252</sup> (Gaus 2012a, 403); emphasis added.

<sup>253</sup> E.g. (Simmons 1999)

same above account of rule-convergence through interdependent choices.<sup>254</sup> He repeats the claim that the equilibrium social rule is “a real choice of actual people”. Insofar as a state only “protects, sustains, and develops our social morality,” then we are coerced by laws to which we as a society have consented through an “on-going” process of interdependent social choice.<sup>255</sup> Therefore, claims Gaus, his ideal liberal state is subject to the actual consent which Simmons says is necessary to generate legitimacy. In another paper he even writes that we can see the equilibrium rule as the expression of a general will in the Rousseauvian sense: “[it is] this social process, which can lead to a uniquely justified outcome but which itself need not be justified [i.e. convergence on a rule through iterative interactions], that allows society to form a general will on a variety of matters without resort to the political process”.<sup>256</sup> Now, Gaus’ arguments here are stylised and refer to convergence on one rule only. If they are successful, however, they can be generalised easily to systems of institutional rules. We would then be able to say that liberal citizens have consented to such a system if all the rules comprising it arose from their interdependent choices.

It would also then seem that there is no necessary tension between an associational understanding of political society which conceives of citizens as exercising collective agency over the institutional rules that govern them and the emergent understanding that sees such rules as the result of human action but no overall design. For on Gaus’ score participation in an emergent process leading to the emergence of an equilibrium rule counts as consent to that rule. Further, to the extent that citizens participate in this emergent process *together* we can say that they have *all* consented to the equilibrium rule. The rule is thus the expression of collective agency on the part of all citizens despite it emerging in an unintended manner. Pace the contention of this chapter and this thesis more generally, the Rawlsian and associational desideratum that liberal citizens create and sustain the institutions that govern them does not fall foul of the Hayekian dictum that large-scale rules are hard to intentionally design. Indeed, Gaus’ claim that we have *consented* to the emergent rule is striking given Rawls’ own reluctance to use the language of consent in his political theory.<sup>257</sup> In assessing Gaus’ claim here we must answer the following question. Do I really consent to a rule if I

---

<sup>254</sup> (Gaus 2012b, 116–20)

<sup>255</sup> (Gaus 2012b, 121–23)

<sup>256</sup> (Gaus 2011b, 91)

<sup>257</sup> While Rawls thinks the well-ordered society is just because of the deliberate intention of liberal citizens to give justice to one another, he also thinks that the exit costs of leaving such a society are too high for us to be able to say they have consented to being a member of it. See, for instance, Rawls (2001, 182).

contributed to the social process leading to its unintended emergence as the equilibrium rule? The next section will reply in the negative.

## 6. Assessing Gaus' Argument

Consider the following intuitive and common sense definition of consent once offered by John Simmons: "First, consent must be given intentionally and (perhaps this is redundant) knowingly; as with promising, one can consent insincerely, but not unintentionally. Second, consent must be given voluntarily."<sup>258</sup> The first part of Simmons' definition captures how intentionality is built into our understanding of consent and indeed agency more generally. It is very hard to see how consenting to be under a state of affairs cannot also entail intending to be under that state of affairs. Hence, we normally say that consent has to be *informed* because one has to know beforehand what they are getting into. Intentionality is built also into our wider understanding of autonomy. Autonomous persons form intentions as to which objectives they would like to achieve and act upon them. Kantian self-legislation just is *intending* to put oneself under the authority of a universal law. So if we are to say that a member of a free society "chooses" or "consents" to the moral rule under which they are governed, then we are also to say that they did so with intention. *Ex hypothesis* however, this condition is not met in emergent social process. They arise through many human actions but no intentional human design. Therefore, Gaus' ascription of consent and choice to the outcomes of this process seems misplaced.

One obvious response from Gaus is to say that the rule in equilibrium has, in fact, been chosen *with intention* by all members of the group, at least *eventually*. Even if they *started* the Kantian coordination game by playing a different rule, each must have eventually switched to the rule which is now in equilibrium. Therefore, they did actually consent (intend) to be under the jurisdiction of that rule from the point at which they started complying with it. One difficulty with this response is made clear by Gaus' own observation that once the rule is in equilibrium it can still bind "one who holds out, refusing to choose our rule".<sup>259</sup> Clearly, it cannot be said of these hold outs who *refuse to comply at all* that they

---

<sup>258</sup> (Simmons 1976, 276); Simmons, correctly in my estimation, says we cannot be very precise about what voluntariness entails save to say that the consequences of saying "no" cannot be excessively deleterious. The traveller who intentionally pays Hobbes' highwayman after being told "your money or your life" has not, then, properly consented to handing over their money.

<sup>259</sup> (Gaus 2012a, 403)

have intended to conform to the rule. Be this as it may, there is a bigger difficulty. The attribution of consent to those who subscribe to the rule *only after* it is established or otherwise close to being established as the equilibrium rule of the community is problematic for a familiar reason.

Suppose I find myself in a community where everyone is expected to conform to rule X. We can posit that the community converged upon the rule by playing Gaus' Kantian coordination game. I am a member of this community and I think very little of rule X. I initially tried to promote the use of other rules but they lost out in the emergent process that led to it becoming the equilibrium rule. Since the rule is now entrenched in the community, my effective choice now is conforming or leaving the community. I have lived in this community all my life and know of little else. I lack financial means to leave. Thus I conform to rule X. Do we conclude that I have *consented* to the rule? The answer seems to be a negative one. The exit costs are so high that we find it difficult to say I *voluntarily* subscribe to X. Thus even if we concede that the emergent social process will meet Simmons' intentionality requirement, at least eventually, his second condition of voluntariness remains unmet. Under these circumstances, it seems much more apt to say I *acquiesce* to X. We might imagine richer and/or more cosmopolitan members of such a community for whom the exit costs might not be as prohibitive as for myself. Still, any society will contain at least some individuals for whom the exit costs are high enough such that we cannot impute consent from their bare compliance. Hume's critique of the social contract hinged on this very point.<sup>260</sup>

Gaus might advance a second response at this juncture. The criticisms advanced thus far misunderstand the role that the emergent process of the Kantian coordination game plays in his argument. We turn to it to solve the indeterminacy that stems from Gaus' deliberative model that comes earlier in the sequencing of his argument. The emergent process of the Kantian coordination game selects only from those rules that are *already* within the socially eligible set. Hence, they are already subject to some form of prior endorsement by all members of the public. One should not be in a situation, such as that of my hypothetical self with respect to rule X, of either subscribing to a rule one cannot endorse or leaving the

---

<sup>260</sup> "Can we seriously say, that a poor peasant or artizan has a free choice to leave his country, when he knows no foreign language or manners, and lives from day to day, by the small wages which he acquires? We may as well assert, that a man, by remaining in a vessel, freely consents to the dominion of the master; though he was carried on board while asleep, and must leap into the ocean, and perish, the moment he leaves her." (Hume 1777d)

community in question. There are two difficulties with this response. Firstly, recall that for a rule to be in the socially eligible set, it simply has to be ranked better than no rule by all members of the public with respect to the issue it covers. However, an improvement on no rule is often a very low bar to meet. I might “endorse” highly authoritarian rules in the sense of ranking them above Hobbesian anarchy, though the former might still fall far short of my first-best rules. That a rule is in the socially eligible set could mean it is subject to no stronger form of endorsement by members of the public than this. Again, it looks like Gaus is presenting what amounts only to Humean acquiescence as a more substantive form of endorsement.

Secondly, and returning to our argument from the previous chapter, in an impersonal society characterised by mutual ignorance, we would be ignorant of the content of the socially eligible set because we do not know what most of our fellow citizens believe. Thus we do not know whether any converged-upon rule is in the socially eligible set. Now, I believe Gaus is at least implicitly aware of this difficulty. His aforementioned pruning strategies seem designed to overcome them. Recall his claim that all persons *qua self-directed agents* would affirm abstract rights of agency and jurisdictional rights. By extension, they would have reasons to endorse rules constituting reasonable interpretations of such rights. We can then *infer* such rules to be in the socially eligible set despite our not having direct knowledge of the concrete views of the mass of our fellow citizens. For we would know that as abstract agents they would have reasons to endorse such rules, to coin a familiar Rawlsian phrase, whatever else they want or believe. The problem, however, is that it does not ineluctably follow that all self-directed agents would affirm the abstract rights of agency and jurisdiction that Gaus claims they would.

To take one already mentioned example, Gaus believes that all such agents would affirm the “presumption in favor of liberty” and by extension any rule that is a reasonable interpretation thereof. This presumption is that individuals are normally at liberty to do as they please such that the burden of justification falls on those who propose to coerce them. Yet, this presumption is contested philosophically. Jürgen Habermas, to name just one high profile philosopher, rejects it.<sup>261</sup> It also seems to be contested by real members of the public. When Gaus himself, along with Shaun Nichols, tested the principle through empirical

---

<sup>261</sup> Habermas claims that the liberties of the moderns and ancients are “co-original”. On his score, public and private autonomy are “not possible without the other, but neither sets limits on the other”. (Habermas 2001, 767)

experiments, they found that “rule followers sometimes suppose a principle of natural liberty [i.e. a prime facie presumption in favour of freedom], sometimes its denial”.<sup>262</sup> So even if we accept that all of our fellow citizens are self-directed agents, we cannot infer *ipso facto* that they all affirm the presumption of liberty and by extension that reasonable rule-interpretations of that principle are in the socially eligible set.

Another example is Gaus’ view that all self-directed agents will endorse a scheme of jurisdictional property rights that are strong in two senses: “[they] are not easily overridden and [which] are extensive, including private property in capital assets”.<sup>263</sup> Again, this does not seem true of all (or perhaps even most) political philosophers. Gerald Cohen’s *Why Not Socialism* is ultimately a paean to the benefits that self-directed agents derive from society-wide collective decision-making.<sup>264</sup> As before then, even if we accept that all of our fellow citizens are self-directed agents, we cannot infer *ipso facto* that they affirm strong and extensive jurisdictional rights and by extension that reasonable rule-interpretations of such rights are in the socially eligible set. Thus, Gaus’ pruning strategies fail to overcome our ignorance of the contents of the socially eligible set. In summary of this section, the most we can say about agents whose actions were causally responsible for the emergence of an equilibrium rule in the Kantian coordination game is that they have acquiesced to that rule. Gaus ascription of consent is inappropriate because the exit costs of leaving the typical nation-state are such that the voluntariness condition necessary to consent is not met. His potential response that the equilibrium rule is subject to a form of prior endorsement in virtue of being in the socially eligible set fails because we do not know its content. Even if we did, the rules that compose it are subject to a very weak form of endorsement i.e. they are merely ranked as better than no rule. For better or for worse, the emergent rules of social morality are something that happen to us. They are not the product of our will.

## Conclusion

---

<sup>262</sup> (Gaus and Nichols 2017, 80)

<sup>263</sup> (Gaus 2012a, 377)

<sup>264</sup> (Cohen 2009)

The first chapter distinguished between the Rawlsian associational understanding of political society and the Hayekian emergent understanding. This chapter has explained why this thesis has resisted the temptation to reconcile these contrasting visions. The task of outlining the theoretical framework is complete. We now turn to the second part of the thesis and the task of vindicating the emergent understanding over the associational one. We will consider and critique the prospects for collective political agency in the “non-ideal” circumstances of actually-existing liberal democracies and the more idealised circumstances of Rawls’ well-ordered society as well as those posited by deliberative democratic theorists.

## **Part II: The Prospects for Collective Political Agency**



## Chapter Three: Descriptive Associationalism

### Abstract

This chapter considers the claim that citizens of actually-existing liberal democracies work together to achieve shared objectives. It examines two models: the *joint-intentional action* model and the *principal-agent* model. The first holds that the citizen-body as a whole acts together. The second holds that the citizen-body as principal appoints an agent in a democratically elected government to act on its behalf. This chapter argues that neither are plausible since both overlook epistemic constraints facing liberal citizens. It also clarifies the implications that the arguments of this thesis have for our understanding of the value of democracy.

### Introduction

Descriptive associationalists hold that citizens of actually-existing liberal democracies are capable of forming and acting upon shared objectives and that the major institutions under which they live are the product of such shared and intentional coordination. As we saw in chapter one, this view is espoused by figures such as Anna Stilz, David Miller, Ryan Pevnick and Margaret Moore. To recap: Stilz conceives of citizens as active “makers” of the institutions under which they live;<sup>265</sup> Miller writes that [nations] “are communities that do things together, take decisions, achieve results, and so forth;”<sup>266</sup> Pevnick sees liberal democratic nations as acting together to create public goods through “the coordinated decisions, labor, and contributions of members;”<sup>267</sup> and Moore claims that actually existing nations have territorial rights because “a large majority [of their] people are in a relationship with one another which is characterized by a shared political commitment to establish rules and practices of self-determination”.<sup>268</sup> What unites authors such as these is the belief that actually-existing liberal democracies are associations through which citizens *do things* together as part of a common enterprise. They fight wars, develop their economy, decide their own laws, determine the nature and extent of their dealings with other states, and so on. This chapter assesses and critiques this characterisation of the actually-existing relationship between liberal citizens.

---

<sup>265</sup> (Stilz 2015, 7–8)

<sup>266</sup> (Miller 1997, 24)

<sup>267</sup> (Pevnick 2014, 38)

<sup>268</sup> (Moore 2015, 34)

It considers two ways in which descriptive associationalism might be theoretically fleshed out. The *joint-intentional* model states that the citizen-body *as a whole* form and act upon shared intentions. This view is explicitly defended by Anna Stilz and Margaret Moore. Both draw on the concept of a joint-intentional action as it has been philosophically defined. Scholars in this field have explicated the conditions under which a grouping of persons can appropriately be described as acting upon jointly shared intentions. The *principal-agent* model posits instead that the citizen-body appoints an agent – a democratic government – who acts on their behalf as their representative. This chapter argues that both models posit implausible knowledge requirements on the part of liberal citizens. The argument to follow is divided into nine sections. Sections 2-4 deal with the joint-intentional model. Section two explicates the necessary conditions for a joint-intentional action, drawing upon the account of Michael Bratman. It highlights the high degree of mutual knowledge required of a grouping of persons party to such action. Section three presents a general argument as to why large-scale citizen groupings lack such knowledge. The claim to the contrary has been developed to the furthest extent by Anna Stilz. Section four thus rebuts her particular argument.

Sections 5-8 deal with the principal-agent model. Section five fleshes out this view in greater detail before subjecting it to critique. Section six argues that voters can express their “will” to governments in only a limited sense. The expression of their preferences is limited to issuing binary judgements on propositions preselected by the political class. The tacit and inarticulate component of voter knowledge is not included. Even if one thought that binary judgements on propositional statements is an adequate way for citizens to express their will, sections 7 and 8 argue that the principal-agent model would still not be plausible. Section 7 argues that while the electorate might issue approval of the stated aim of a policy, its substantive implementation is always at the autonomous discretion of state officials. Perhaps the discretionary authority of state officials would not be such an issue for the principal-agent model if citizens could *retroactively* judge governmental performance. However, section eight argues that citizens also lack the requisite knowledge to undertake such evaluation. On the basis of these epistemic considerations, it is concluded that the state cannot be understood as the agent of the citizenry. Finally, section 9 discusses the relation of this chapter’s argument and that of the thesis as a whole to claims as to the value of democracy. While this thesis is sceptical with regard to intrinsic claims as to democracy’s value, it is quite compatible with instrumental and even epistemic such claims. In this chapter we take into account how compliance problems, stemming both from ethical disagreement

and mundane self-interest, effect the prospects for collective political agency in addition to epistemic problems. We thus relax the strict compliance assumption adopted elsewhere in the thesis. This seems appropriate as we are here considering actually-existing citizens and state actors with their actually-existing motivations and beliefs.

### 1. The Joint-Intentional Action Model

What do we mean when we say that a group of persons does something together? A common sense answer would be that together they formulate a shared goal and try to achieve it. There is by now a well-established philosophical literature formally explicating the conditions that need to hold for a group of persons to be able to act together in such a way. At first approximation, we can say that both the intention to be realised and the means by which it is to be realised need to be common knowledge between the members of the group if they are to relevantly share the intention in question.<sup>269</sup> It is only natural then that political philosophers look to this literature when they try to explain how a group of citizens might act together. Margaret Moore defines a “people” as a collective agent that is the appropriate holder of territorial rights. A collective agent, in turn, refers “to an entity with two or more members that can perform joint actions” as they are technically understood in the philosophical literature.<sup>270</sup> Anna Stilz has in various works also argued that citizens are capable of joint-intentional actions as that concept has been philosophically explicated. Explicitly building on Michael Bratman’s account, she claims that in obeying the law, dutifully paying their taxes and exercising their say in democratic politics, liberal citizens all do their part in the joint-intentional activity of realising a common scheme of just law.<sup>271</sup> This section explains the philosophical puzzle that the joint-intentional action tries to solve. Additionally, it will detail Bratman’s account of the demanding degree of mutual knowledge required of parties to a joint-intentional action. Sections 2 and 3 will explain why these conditions are not met in the typical citizen grouping.

The philosophical literature on joint-intentional actions arose to solve a puzzle that arises when one tries to make sense of the idea of several agents doing something together. Suppose two people decide to go for a walk. On their way out, a third person asks them what they are doing. One replies, “we have decided to go for a walk”. The implication is that they

---

<sup>269</sup> (Bratman 1999); (Gilbert 1992); (Searle 1990); (Tuomela 2005)

<sup>270</sup> (Moore 2015, 21)

<sup>271</sup> (Stilz 2009, chap. 7); (Stilz 2015, 105–16); (Stilz 2016, 13–14)

both have a shared intention to embark upon the joint-activity. However, if an “intention is a mental state or event, and minds belong to individual persons” then it seems unclear how more than one person can share the same intention.<sup>272</sup> One seemingly easy solution is what we might call the *summation strategy*. This would hold that strictly speaking there is no “shared” intention. One says “we have decided to go for a walk” as a convenient short-hand for something like the following: “she and I have independently decided to go for a walk at the same point in time for the same duration of time and over the same route”. In other words, the phenomenon is really just a summation of separate individual intentions that coincide.

This is conceivably true of some forms of joint-activity. Consider a group of passengers engaged in the joint-activity of riding an underground train. Suppose I, perhaps in the guise of an eccentric anthropologist, walk onto a London Underground train and ask one of the passengers “what are you doing?” One of the (probably puzzled) passengers might well decide to speak for both herself and the other passengers on the train and reply “we are riding the train”. Clearly, however, in this instance she is just summarising the distinct and independently-formed intentions of each passenger to ride the train in order to arrive at a particular destination. The destinations might well coincide, as on a hot day when many passengers might separately decide to visit the park. Still, an intelligible account of each passenger’s intentions does not require reference to the intentions of the other passengers. One passenger’s decision to visit the park on a hot day is intelligible *regardless* of why the other passengers happen to be on the same train. However, the summation strategy can’t account for cases where agents’ intentions are clearly coordinated. Consider a group of musicians playing a song. In this instance, we cannot give an intelligible account of each person’s actions without reference to the intentions of the other band members. Here we do have to offer an account of how intentions are mental states of discrete individuals and yet shared between them. This is the goal of the joint-intentional action literature.

Michael Bratman’s theory is one of the leading such accounts. On his account, one cannot intelligibly say “I intend to J” if my J-ing is a joint-performance with other people. I can only intend to perform those actions that are solely my own. However, we do speak of “intentions-*that*” in reference to our intentions to influence the actions of others – I intend *that* my child cleans up their room.<sup>273</sup> It follows that I can speak of “my intention that we J”

---

<sup>272</sup> (Velleman 1997, 29)

<sup>273</sup> (Bratman 1999, 142)

despite my not being the sole author of that act. Assuming a group consisting of you and I, we intend to *J iff*

1. I intend that we *J* and I intend to do my part in our *J*-ing if I know that you intend that we *J* and intend to do your part in our *J*-ing *and*
2. I intend that the sub-plan according to which I will carry out my part *meshes* with your equivalent sub-plan.
3. The equivalent is true of you.
4. This is common knowledge between us.<sup>274</sup>

The first clause renders my intention to perform my part of the joint-performance conditional on knowledge of your reciprocation. It is important to note that this is not a one-off signal between the parties. In another formulation of the above, Bratman says that the “persistence” of my intention that we *J* depends on my “continued” knowledge of your reciprocal intention over the course of the activity.<sup>275</sup> The second clause is of great importance. For Bratman it is insufficient that all parties continuously share – and per clause 4 know that they all share - an *abstract* commitment to *J*.<sup>276</sup> It also needs to be common knowledge between them that the sub-plans with which they intend to fulfil their parts in the collective act are compatible. In Bratman’s terminology, they need to *mesh*.

Specifically, the parties need to have shared understandings as to how they (a) coordinate their intentional actions; (b) coordinate their planning; and (c) structure relevant bargaining. He illustrates this with the example of two people painting a house. They need to coordinate their actions by, for example, ensuring that the scraping is performed before the new paint is applied. They need to coordinate their planning such that if one supplies the paints but not the brushes, the other knows to bring the brushes. They also need relevant bargaining methods to settle any disagreements such as when they disagree on which colour to use.<sup>277</sup> While Bratman does not specify this, the coordination of sub-plans would also need to be on a continuous basis. Thus we can add our own fourth desideratum. The parties need be able to (d) adjust to changing circumstances since not everything can be planned for or predicted

---

<sup>274</sup> (Bratman 1999, 121). I take the liberty here of modifying Bratman’s own formulation of the necessary and sufficient conditions. I hope this rendering, while still a fundamentally accurate interpretation, reads somewhat more intuitively if not quite as technically precise as Bratman’s original rendering.

<sup>275</sup> (Bratman 1999, 153)

<sup>276</sup> The fourth clause is designed to avoid a familiar problem of iteration. It is insufficient that I know that you intend to do your part and that our plans mesh. *You* need to know that I know *this*. In turn, I need to know that you know *that*. And so on ad infinitum.

<sup>277</sup> (Bratman 1999, 112)

in advance. If on the day our painters find they will need more paint than they anticipated or that impending rain has thrown a spanner in the works or that the job specifications change, they will need to re-coordinate, as it were, to a new settlement on how they are to proceed. To summarise, members of a group share an intention that they J *iff* each has *continuous* assurance that all others intend that they J and that their contribution is compatible with the contributions of others. When these conditions are in place, we can genuinely speak of an irreducibly shared intention rather than the ultimately individual intentions of the summation strategy.

## 2. The Infeasibility of Joint-Intentional Action Under Large-Scale Settings

It is evident from our review of it that the mutual knowledge requirements of Bratman's model are quite demanding. They will be met only in small groups in which the parties have a high degree of proximate knowledge of one another. It is telling that Bratman's own examples are of a handful of people performing very closely coordinated activities such as painting or singing a duet.<sup>278</sup> Nor is he the exception in the literature. John Searle uses the examples of a football team and an orchestra.<sup>279</sup> Margaret Gilbert refers to two people going on a walk.<sup>280</sup> Tuomela speaks of a group of persons who clean a park.<sup>281</sup> There are quite straightforward reasons to think that as the number of agents increase and their circumstances become more diverse, the capacity of a grouping to undertake joint-intentional actions diminishes. If the counterparties are a handful of agents, I can obtain quite detailed knowledge of their willingness to participate in the group activity and how their sub-plans might mesh with mine. And vice-versa. As the group gets larger, this knowledge becomes ever more difficult to obtain. When the grouping-in-question is the number of individuals who comprise the typical liberal democratic nation-state, these mutual knowledge conditions are not met. One is barely aware of what intentions they might share with millions of fellow citizens. Even if we posit that all citizens share some abstract intention, remaining would be the byzantine task of ensuring that each individual's proposed

---

<sup>278</sup> (Bratman 1999, 99, 103)

<sup>279</sup> (Searle 1990, 413)

<sup>280</sup> (Gilbert 1992, 362)

<sup>281</sup> (Tuomela 2005, 336)

contribution to the shared objective – their “sub-plan” in Bratman’s terminology – is suitably coordinated with those of all the others.<sup>282</sup>

This returns us to the key claim of the emergence view. Joint-action in Bratman’s sense requires agents who have personal knowledge of one another. These knowledge conditions do not hold under impersonal settings. The function of liberal institutions is to facilitate coordination under such conditions where joint-intentional action is not feasible. Such coordination leads to emergent outcomes because it is composed of the interactions of very many agents acting upon their own idiosyncratic aims and knowledge. On the Hayekian score, to apply Bratman’s model to impersonal relations would constitute a category mistake. If we insist on saying that actors involved in impersonal coordination are undertaking “collective action,” we should do so only in the weak sense of the summation strategy. We shall look at her arguments in more detail below, but for now let us consider a specific objection Anna Stilz would raise against the claim we have just made. It is that there seem to be groupings that are large enough to be impersonal but which might still plausibly be described as capable of a joint-action. She writes:

Even when an institution is too large for all its members to know one another, it can still feature significant levels of cooperation. Though I do not know most of the other students and employees at my university, still I “play my part” in it willingly. Many large groups—think of megachurches, or trade unions—do not exhibit close personal ties among the participants, yet their activity nonetheless counts as shared cooperation.<sup>283</sup>

Now, whether cooperation is “shared” or “significant” is not the issue here. Impersonal and very large coordination on the market can also be so described. What is at stake is whether the cooperation is on the basis of jointly shared intentions. I take Stilz to argue that such is also the case here. Despite not knowing one another personally, the university students and employees can each say “I play my part” in a shared endeavour. There does seem to be a sense in which trade unions or universities can share participatory intentions towards shared goals despite their impersonal relations. How else do we account for the idea of strike action for instance? Many impersonally related union members manage to assure one another of their shared commitment to strike and indeed render compatible their sub-plans by, among other things, turning up to the picket line on the same day. If so, and pace our analysis above,

---

<sup>282</sup> Of course, Hayek’s point was that prices tacitly facilitate this kind of coordination within the context of a market economy. It cannot be facilitated *consciously*.

<sup>283</sup> (Stilz 2016, 107–8)

it seems that impersonal relations do not pose any categorical impediment to joint-intentional action. If it is possible between thousands of impersonally related university members then it should also be possible between millions of impersonally related citizens.

One difficulty with this objection is that it is far from obvious that an organisation such as a university indeed has a “shared goal” in relation to which all staff and students can say “I play my part”. Those who do think there is such a goal would be tempted to refer to a lofty academic ambition such as the dissemination of knowledge. Yet this is really true only of a subset of those that associate with the university. Many undergraduate and graduate students associate with the university as a means *to* gainful employment. Non-academic staff associate with the university as a means *of* gainful employment. It is arguably only permanent academic staff who associate with the university as a means to further a lofty academic goal. Even here matters are more fractured than they seem, since academics will have different understandings of what their “shared” academic goal might be.<sup>284</sup> So one could push back against Stiliz’s objection by saying that the members of a university are really closer to our train passengers in their relations than Bratman’s painters.

In any case, even if we accept the premise that university members can and do coordinate their actions towards shared objectives, it does not necessarily follow that the same can be said of a group consisting of millions of liberal citizens. To see why, let us draw an analogy with the paradox of the heap. One or two grains of sand do not constitute a heap. A million do. We cannot specify the minimum number of grains necessary for a bona fide heap. Still this does not prevent us from identifying clear-cut cases on either side of this indeterminate threshold. We can think of the knowledge necessary for joint-action as a vague predicate in this way. We know for sure that a handful of persons do have sufficient mutual knowledge to undertake a joint-intentional action. We could also accept that some impersonal groupings of an intermediate size also have sufficient such knowledge. Let us grant that this is true of most universities. As the grouping-in-question grows in size and complexity, however, a point will be reached at which it no longer has the requisite knowledge. Quite where this indeterminate threshold might be does not matter for our purposes. The question is whether the citizen grouping of the typical nation-state is of such size and complexity to

---

<sup>284</sup> These might range from discovering the truth to disseminating the truth to improving public policy to furthering social justice and so on. Even if we restrict ourselves to one such academic goal, they have different meanings in different disciplines. Consider what “discovering the truth” might entail in physics as opposed to political philosophy.



be definitively above this threshold such that we can also definitively say that they lack the knowledge necessary for joint-intentional action.<sup>285</sup> What reasons might there be to answer this question in the affirmative? Three might be advanced.

First, members of a university have chosen to join the university *and* we can identify at least some putative aims to which a university might be dedicated to upholding.<sup>286</sup> It might then seem plausible to infer that the members of that university share the intention to uphold these putative aims. Neither condition seems to hold with respect to political associations. Most citizens have not opted to join the political associations of which they are a member. Nor can we identify even any putative aims uniting citizens of diverse countries such as the United Kingdom or United States.<sup>287</sup> Second, we know from economic theory and practice that the knowledge necessary for conscious and centrally coordinated planning can be ascertained in the intermediate sized grouping of the corporation. We also know it cannot be ascertained by central planners attempting to direct the economy of whole societies. This fact alone suggests there is a qualitative difference between the ease with which information can flow within groupings of the size and complexity of the typical university or company compared with groupings the size of a nation state.

The third reason is probably the most important and consequently deserves some elaboration. The oldest universities in the world do seem to exhibit continuity and unity of purpose over very long periods of time. Universities such as Oxford, Cambridge, Bologna, Padles or Al-Karaouine have been engaged in some form of scholarly activity for centuries and in some cases over a millennium. We are often tempted to ascribe this same continuity and unity of purpose to political associations. However, large-scale political institutions do not arise as the shared project of the liberal citizenry. They are indeed better described as emergent phenomena. They are the unintended result of strategic interactions between groups acting on the basis of their own knowledge and values. This does not mean that the

---

<sup>285</sup> One might wonder what is meant by a “typical” state given that the smallest state in the world, Vatican City, has less than a 1,000 residents while the largest state in the world, China, has nearly 1.4 billion residents. By “typical” I mean states whose populations run at least into the millions. This is true of the majority of states in the world. Additionally, the vast majority of the world’s population live in such states. By way of some comparison, the London School of Economics has just under 13,000 staff and students. Very small states such as those found in the Pacific Ocean or those that are better described as principalities, such as Monaco or Liechtenstein, might well be exempt from the argument offered here. At the same time, however, advocates of the associational model hardly restrict the application of their theories to such small states.

<sup>286</sup> Such as those identified above.

<sup>287</sup> Rawlsians might be tempted to refer to social justice as a shared aim. I respond to this in the section below.

actions of all agents are equally efficacious. We can expect some actors – e.g. political elites or those who are especially effective at political mobilisation – to have a greater impact on events. Still, the course of events often outrun even their intentions.

Several considerations can be offered in support of this contention. Even the most paradigmatic examples of collective action for political ends, mass protests leading to revolutions, tend to involve at most a small minority of the population. The French Revolution of 1789, the Russian Revolution of 1917 and the Romanian Revolution of 1989 are estimated to have involved no more than one per cent of the population at the time. The 1979 Iranian Revolution is claimed to have had the biggest turn-out but it was not more than ten per cent of the population.<sup>288</sup> Further, those who do take part in these events do not always see themselves as taking part in a “Revolution”. Our terminology here is potentially misleading. The characterisation of these events is *ex-post*. The actors who took part in these events on-the-ground may not have seen themselves as contributing to a seismic historical event. As James Scott observes:

Those who stormed the Bastille to free prisoners and seize arms could not possibly have known (much less intended) that they would bring down the monarchy and aristocracy, let alone that they were participating in what later would come to be known as “the French Revolution.”<sup>289</sup>

Nor do the disparate groups who take part in those events subsequently designated as “Revolutions” share the same objectives other than perhaps wanting to change the status quo. The Iranian Revolution, for instance, consisted of various groups ranging from secularists to Marxists to religious inspired groups. The religious groups themselves were heterogeneous, with some wanting to keep religion out of politics, others fusing their Islamism with Marxism and yet others following the doctrine of the figurehead of the Revolution, Ruhollah Khomeini.<sup>290</sup> It is very safe to say that not all members of the revolutionary coalition intended the regime that Khomeini would subsequently set up after he and his faction won the post-revolutionary power struggle.

Similarly, a historical survey of major constitutional moments, reveals that most of the population have not been involved. Rather these were elite-led projects in which the masses

---

<sup>288</sup> These figures were obtained from Kurzman (2005, 122–23).

<sup>289</sup> (Scott 2012, 137)

<sup>290</sup> On this see Kurzman (2005).

were swept up. Hume famously wrote of the Glorious Revolution in Britain – “And it was only the majority of seven hundred, who determined that change for near ten millions”.<sup>291</sup> Russell Hardin notes that the deliberations of the Constitutional Convention of 1787, out of which emerged the US constitution, were “essentially” a secret and remained so until the death of the conventioners.<sup>292</sup> When the French National Constituent Assembly issued the Declaration of the Rights of Man and of the Citizen in 1789, the majority of French “citizens” didn’t speak French and were primarily loyal to their local village rather than the French state.<sup>293</sup> None of this should be taken to imply that elites are puppet masters who manipulate and control events from behind the scenes as they wish. For one thing, “political elites” are typically a heterogeneous category of actors with distinct interests and aims. For another, political economy and political sociology suggest that not even elites planned or intended the rise of the contemporary state as we know it. It is best seen as the unintended by-product of strategic interactions occurring primarily between different elite groupings.

This is the account offered by North, Wallis and Weingast in their documentation of the transition from “natural states” to “open access orders”. They use the latter term to describe the institutions of liberal democratic states. The hallmarks of such orders is *impersonal* governance, whereby rights and entitlements are distributed to citizens in a standardised and impartial manner. The aspect of open-access orders that are most important to the authors is universal access to perpetually lived organisations.<sup>294</sup> The hallmark of natural states is governance on the basis of personal relationships – “who one is and who one knows, form the basis for social organization”.<sup>295</sup> They designate natural states as limited access orders because the right to form organisations under them is limited to well-connected elites. These actors form a coalition who divide up privileges and rents between one another while excluding outsiders.<sup>296</sup> They note that such elites “are not a unified group, but are

---

<sup>291</sup> (Hume 1777d)

<sup>292</sup> (Hardin 1999, 111); further the manner in which the US Constitution was ratified owed more to network and coordination effects than any joint-intentional action of US citizens. We shall return to this point in section five of chapter seven.

<sup>293</sup> These “peasants” would be turned into “Frenchmen” over the nineteenth century, as Eugen Weber (2007) famously documented.

<sup>294</sup> (North, Wallis, and Weingast 2009, 21–25); I refer here to *Violence and Social Orders*. There is some overlap between their understanding of an open-access order governed by impersonal law and Hayek’s understanding of the Great Society governed by purpose-independent rules of just conduct.

<sup>295</sup> (North, Wallis, and Weingast 2009, 2)

<sup>296</sup> (North, Wallis, and Weingast 2009, 18–21)

composed of disparate groups that compete and cooperate, and sometimes go to war against each other”.<sup>297</sup>

They seek to uncover and explain the processes by which societies transition from natural states to liberal democratic open access orders. Using the historical case studies of France, the United States and Britain, they offer the following explanation. In the 18<sup>th</sup> century, many members of the elite coalition had the incentive to convert their privileges into more formal legal rights because they feared that “intra-elite political competition would inevitably lead to consolidated political control by a faction that maintained control through manipulation of the economy”.<sup>298</sup> In other words, some groups in the elite coalition preferred to turn their privileges into more legally robust “rights” in order to prevent predation from other groups in the elite coalition. The creation of one set of elite rights created incentives for elite-actors to credibly extend such rights yet further:

The privilege of owning shares in a joint-stock company, for example, may begin as a unique privilege. However, if that privilege becomes widespread and shares are transferable, then an elite interest in supporting impersonal exchange of shares may grow. Impersonal exchange in shares, in turn, may create an interest in impersonal formation of companies, forces that clearly came to the fore in all three countries in the nineteenth century.<sup>299</sup>

The crucial point here being that these actors did not set out to create open-access and liberal political institutions. At each juncture, they were responding to incentives out of self-interest. The emergence of liberal institutions was the cumulative and unintended effect of these changes.

In his own famous account, Charles Tilly sees contemporary state institutions as the unintended outcome of the efforts of political elites to create a fiscal base for war. Tilly emphasises that states compose a *system* “to the extent that they interact, and to the degree that their interaction significantly affects each parties fate”.<sup>300</sup> Within this system, state institutions were more effective than alternative institutions in gathering manpower to fight wars and accumulating debt finance to fund them. The political elites who had such modes of political organisation at their disposal had a military advantage over those who did not.

---

<sup>297</sup> (North, Wallis, and Weingast 2009, 149)

<sup>298</sup> (North, Wallis, and Weingast 2009, 256)

<sup>299</sup> (North, Wallis, and Weingast 2009, 256)

<sup>300</sup> (Tilly 1992, 4)

State institutions as we know them arose through this evolutionary selection affect.<sup>301</sup> The political elites did not anticipate this as the outcome of their actions. The state is better seen as the unintended result of a “mosaic of adaptations and improvisations”.<sup>302</sup> Cardinal Richelieu, for example, did not intend to lay the foundations for the “French State” when he sent central government officials to collect tax funds rather than relying on local feudal authorities. He did so because the French crown was in desperate need of war finance and those local authorities had stopped cooperating financially.<sup>303</sup>

Hendrik Spruyt also offers an evolutionary account similar in many respects to Tilly. For Spruyt there is no inevitable or necessary transition from the dispersed and polycentric mode of political authority in the Middle Ages to the contemporary state. The emerging Weberian state faced competition from other political forms, such as city states and city leagues.<sup>304</sup> Spruyt emphasises the superior organisational logic of the Weberian state as the selection mechanism rather than its sheer capacity for resource mobilisation. Such states were economically more efficient and were better placed to make credible commitments to external counterparts. Network affects are also part of Spruyt’s story. Once the number of Weberian states reached critical mass and became the standard mode of political organisation in the international system, other alternatives were de-legitimised.<sup>305</sup> For Spruyt “we cannot simply deduce institutional outcomes from preferences or impute preferences from observed outcomes”.<sup>306</sup> Instead, and like Tilly, he sees states institutions as arising within the context of strategic interactions between elites in an international system.<sup>307</sup> In summary of this section, even if we grant that intermediate groups such as universities are capable of joint-intentional action, there are definitive reasons to think that large citizen groupings are not. Such intermediate associations have at least some putative purposes and their members join them voluntarily. The analogous statement cannot be made of citizen groupings. Economic theory and history suggest a qualitative difference between these kinds of groups in the degree to which information can be shared within them. Finally, as we have just seen, history does reveal a continuity of purpose in the world’s

---

<sup>301</sup> (Tilly 1992, 30–31)

<sup>302</sup> (Tilly 1992, 26)

<sup>303</sup> (Tilly 1992, 26)

<sup>304</sup> (Spruyt 1994, 18–21)

<sup>305</sup> (Spruyt 1994, 28)

<sup>306</sup> (Spruyt 1994, 26)

<sup>307</sup> For a wider review of the sociological literature consistent with this finding, see Cavanaugh (2011, 12–18).

oldest universities but also shows state institutions to be emergent phenomena arising out of elite interactions rather than the creation of a civic-minded citizenry.

#### 4. Liberal Justice as the Object of Joint-Intentional Action

We have completed our general critique of the claim that mass groupings of citizens are capable of joint-intentional action. We now turn to Stilz's specific argument. As already mentioned, hers is the most explicit and elaborate version of this claim. Citing Bratman, Stilz says that I "share a participatory intention whenever I think of myself as doing something because we are doing something together".<sup>308</sup> The state itself, she claims, can be understood as just such joint-intentional activity but on a "grand scale".<sup>309</sup> On her Kantian account, liberal justice requires a unitary system of coercive law to enforce citizens' right of person and property. These need to be "objectively" defined and enforced by the state if we are to avoid the conflicts that arise when private parties try to "subjectively" define and enforce their rights.<sup>310</sup> The objective definition and enforcement of the law is the object of the joint-intentional agency of liberal citizens. For laws cannot operate solely on the basis of state coercion,

... they depend much more pervasively on large-scale patterns of behavior on the part of people who orient their actions to these laws, who take law to have a moral salience that affects their practical reasoning about what to do. By obeying the law, then, residents of a particular state help to make their acquired rights an affective reality... Likewise, by paying taxes, members actively contribute to sustaining the institutions that enforce those rights against those individuals who refuse to voluntarily comply.<sup>311</sup>

In addition to obeying the law and paying our taxes, we share also the intention of participating in their formulation. Citizens act on this shared intention by voting in addition to "[taking] part in social movements, debating political issues, voting for representatives, and retaining a right of participation in the revision of their constitution".<sup>312</sup> Despite citizens not knowing one another personally, then, their complying with the law and their political activity can be seen as a form of joint-intentional activity with the aim of realising a common and just scheme of law. Stilz's model has much in common with Rawls' thought that just

---

<sup>308</sup> (Stilz 2015, 13)

<sup>309</sup> (Stilz 2015, 13)

<sup>310</sup> (Stilz 2009, chap. 2)

<sup>311</sup> (Stilz 2009, 193)

<sup>312</sup> (Stilz 2009, 194)

political institutions are appropriately secured when citizens act justly with one another rather than have justice imposed on them.

Unlike Rawls, however, Stilz' account is intended as a description of the relationship between actually-existing liberal citizens. It does not refer to an idealised well-ordered society. With this in mind, we can identify three difficulties with her account. First, as argued in chapter two, it is conceptually mistaken to hold that states produce a unitary system of law. State institutions do not speak with one voice because they consist of multiple legislative bodies and agencies. To return to an example from that chapter, citizens of the United States can be subject to the laws of the United States Congress, State legislatures, City governments, municipal governments and county governments. It is implausible to think that the manifold laws and regulations issued by state institutions embody any overall shared intention, such as the upholding of liberal rights. Many of the laws issued by putatively liberal states, such as the prohibition of recreational drugs, are hard to reconcile with any liberal conception of justice. Second, not all state laws are seen as legitimate by the populace and are consequently disregarded by them. Widespread violation of laws banning drug use are but one example of this. Widespread efforts to minimise tax exposure are another. Third, political engagement is at best haphazard in most liberal democracies. Over the last twenty years, voter turnout in the United Kingdom has ranged from 58% to 68%. The turnout for the high profile Brexit Referendum, frequently commented upon as exceptionally high compared with general election turnouts, was 72%. The equivalent variance for United States Presidential elections ranges between 50% and 60%. In sum, Stilz exaggerates the coherence of states' legislative output, the extent of citizens' legal compliance and citizens' commitment to political participation. The biggest difficulty with this account, however, is that it is far from clear what coincidental motivations liberal citizens might have in obeying the law. Surely only few citizens affirm the explicitly Kantian reasons offered by Stilz. Others might have other kinds of philosophical justifications in mind and there is no reason to think that they will be commensurate or compatible with Kantianism. Many will presumably obey for conventional reasons. They will acquiesce to state authority without deeply reflecting on the matter because they have been socialised to do so. Still others will obey the law for the instrumental reason of avoiding state punishment and/or social censure if they are caught.

Many of us will have an ambiguous mixture of these and other conceivable kinds of motivation.<sup>313</sup>

Anticipating this criticism, Stiliz writes:

... It is not necessary that each member who obeys the law and pays taxes *actively intends* the entire system of public coercion that the state puts into effect. The resident may simply be intending not to get himself punished, or to obey the law, and *may not be in any way conscious* of the acts of his compatriots and their overall connection to a system of [just] public coercion. Still, we can say that the residents are in fact contributing to the [just] public coercion of one another even though they don't intend the entire project. Our resident's actions are *causally linked* to an ongoing collective process – the public coercion of other people on a state's territory in the name of justice – and he can be expected to be aware of this fact, *if he reflects on it*.<sup>314</sup>

On the face of it, this passage just seems contradictory. Either the citizens participate in the shared intention or they don't. If Stiliz here concedes that many don't share the relevant intention to realise an objective system of law, then she cannot simultaneously claim they are participants in the joint-intentional action to that end. This said, a more charitable interpretation can be gleaned from the last sentence of that passage. It would run as follows. Were someone to *reflect* on the effects of their legal compliance, they would realise how it is causally linked to the realisation of just laws and – though Stiliz does not explicitly say this – would also *endorse* it. Thus, they can still be described as participating in a joint-intentional action insofar as their behaviour is casually connected to the realisation of an end that *they would endorse upon reflection*. The problem is that this just seems to beg the question. Perhaps after learning about and reflecting upon prevailing political institutions, a previously apathetic person would endorse them. Or perhaps they won't. Perhaps intellectual reflection would even make them more sceptical about politics than they were previously.<sup>315</sup> We cannot tell *a priori*. Thus, Stiliz's description of actually-existing liberal citizens as sharing the objective of realising a unitary and "objective" – in the Kantian sense – scheme of law is

---

<sup>313</sup> The problem here, however, is not simply that our rationales for obeying the law might differ. On Bratman's model, each intends to do their part in the collective action because they *know* the equivalent is true of all others. To act jointly we do not simply need to have the same intentions but each needs to *know* that all the others have the same intentions. Even if we assume full compliance in this scenario – i.e. posit that all citizens affirm Stiliz's Kantian rationale for obeying the law and are motivated to do so – they would still need to *communicate* this fact with one another and such assurance is very hard to come by between large groupings of liberal citizens. We shall develop this argument in the next chapter.

<sup>314</sup> (Stiliz 2009, 193); emphases added.

<sup>315</sup> They might even flirt with anarchism if their intellectual reflection exposes them to philosophers such as Huemer (2013) and Simmons (1999).



mistaken. So concludes our critique of the joint-intentional action model. We now turn to what might initially seem like a more promising model, the principal-agent account.

## 5. The Principal-Agent Model

Perhaps the joint-intentional action model is just excessively demanding in its requirement that *all* members of the group share not only the relevant intention but also assure themselves that everyone else shares it. Surely, it might be objected, a plausible model of collective political agency does not have to meet such stringent criteria. The principal-agent account might be a better candidate. Citizens democratically elect governments. Governments make legislative decisions for the country and have the executive and judicial power to enforce them. Insofar as they are democratically accountable to the electorate the decisions and actions of the government can be ascribed to the citizenry as a whole. For the agent is authorised to act only on behalf of the principal to whom it is ultimately accountable via the ballot box.<sup>316</sup> This model is popular too among liberal associationalists. As we saw earlier, David Miller says that the “policies pursued by the state can reasonably be seen as policies for whose effects the citizen body as a whole is collectively responsible, given that they have authorized the government to act on their behalf in a free election”.<sup>317</sup> Paul Kelly refers to the collective will of citizens as expressed by “democratically legitimate legislative institutions”.<sup>318</sup> Rawls too invokes this model rather than the joint-intentional action model in his international theory. When referring to the “rational conduct” of “peoples,” he says it is “organised and expressed in their elections and votes, and the laws and policies of their government”.<sup>319</sup>

The plausibility of the principal-agent model rests on the ability of voters to communicate their preferences to the governing classes. There are two ways in which they might do this.

---

<sup>316</sup> The following example might help clarify the distinction between the joint-intentional and principal-agent models. The Apollo 11 moon landing could not at all plausibly be described as a joint-intentional action by all American citizens. It was proposed as a national goal by President Kennedy in 1961 and was discharged by a specialised state agency in NASA in 1969. The bulk of the United States’ citizens were involved neither in the formulation nor implementation of this goal. Still, the principal-agent model would hold that it was a “national achievement” of the United States as a whole because it was undertaken by a government duly elected by and acting on behalf of the sovereign citizenry.

<sup>317</sup> (Miller 2007, 128)

<sup>318</sup> (Kelly 2005, 73)

<sup>319</sup> Rawls (2003, 25); again, though, Rawls is referring to idealised rather than actually-existing regimes. We are here considering the plausibility of the principal-agent model as a description of the relationship between actually-existing liberal citizens.

*Ex-ante* accountability – voters select a party or candidate on the basis of their professed policy manifesto or commitments. Voters thus signal their approval of a policy prior to its implementation.

*Ex-post* accountability – voters judge retrospectively the outcomes produced by a governing party. Their continuing to vote them into office signals their retrospective approval of their policies.

The principal-agent model is undoubtedly more plausible than the joint-intentional action model. It takes much more seriously the circumstances of a large-scale society marked by the fragmentation of knowledge and an extensive division of labour. Government would inevitably be a specialist activity under such circumstances rather than a shared endeavour to which all citizens contribute. In focusing on how the governed can hold their rulers to account rather than how all citizens might act-in-unison, the principal-agent model is much more realistic than the joint-intentional alternative. However, we shall argue that it too ultimately fails. This is not to deny that the right to vote provides some mechanism of accountability between the government and the electorate. Democratic accountability, for instance, is famously effective at preventing famine.<sup>320</sup> Effective democratic accountability, however, tends to be possible only when there are “clear, easy-to-recognize policy failures that can readily (and correctly) be ascribed to government officials”.<sup>321</sup> For the most part, however, state and governmental institutions display significant autonomy from the intentions, beliefs and values of individual voters.<sup>322</sup>

Three considerations will be advanced in defence of this claim. First, democratic choice is limited to options pre-selected by the political class and expressed in explicit propositional form. However, the bulk of voters’ preferences and values are tacit and embodied in their respective ways of life. This knowledge cannot be transmitted to the government via the ballot box. Second, even if we posited that voter beliefs were adequately expressed via their votes on propositional statements, we are still faced with the problem of indeterminacy. Abstract policies require interpretation to be implemented and interpretation is at the discretion of politicians. Voters do not authorise the concrete manner in which policies are implemented. Third, knowledge problems make it very difficult for voters to evaluate *ex-post*

---

<sup>320</sup> (Sen 1981)

<sup>321</sup> (Somin 2013, 192)

<sup>322</sup> Here we set aside chapter two’s claim that the state is more like a polycentric network than a unitary organisation. If successful, that argument by itself undermines the principal-agent model. For there would then be no unitary or coherent “agency” capable of “representing” the electorate. The argument presented in this section amounts instead to the following. Even if we did conceive of the state as a unitary agency, the principal-agent model would *still* not be plausible because of the autonomy such an agency would enjoy with respect to voter preferences.

the performance of governments. The first and second considerations are advanced to undermine the claim of ex-ante authorisation. The third undermines the claim of ex-post authorisation.

## 6. Tacit Knowledge and the Limits of Preference Aggregation

We can begin by noting the well-known aggregation problems of social choice theory. Arrow famously showed that in a choice between three or more options undertaken by two or more people, there is no decision procedure that produces a complete and transitive social ranking and does not violate one of several reasonable fairness constraints.<sup>323</sup> Examples of the latter include *universal admissibility of orderings* – individuals may choose from any possible rank ordering of the three options; and *non-dictatorship* – there is no individual who unilaterally decides the social ranking.<sup>324</sup> More recently, Christian List has shown that when three or more people are tasked with issuing “yes/no” answers to a set of logically connected propositions, there is no decision procedure that can produce a logically consistent set of collective answers and meet two fairness constraints. The latter two are *robustness to pluralism* – the procedure accepts as an input any individually rational set of answers to the propositions under consideration; and *basic majoritarianism* – the collective acceptance of any particular proposition on the agenda depends on majority acceptance.<sup>325</sup> In both cases the dilemma is similar. If we are to avoid collectively irrational decisions – the violation of transitivity and completeness criteria for Arrow; the violation of a consistency criterion for List – we are forced to select among an array of decision procedures each of which violates a reasonable constraint of fairness. Any meta-decision procedure for deciding on the choice procedure will itself be subject to these kinds of difficulties. The content of the people’s will is thus arbitrarily contingent on the voting procedure.

Such conventional aggregation problems refer to the difficulties of aggregating individual voter preferences *with respect to binary judgements on pre-selected propositional statements*. That is, voters are restricted to either accepting or rejecting such statements.<sup>326</sup> The great achievement of social choice theory is to show that the public may not speak with

---

<sup>323</sup> (Arrow 1950); I here follow Riker’s (1988, chp. 5) account in *Liberalism Against Populism*.

<sup>324</sup> (Riker 1988, 116–19)

<sup>325</sup> (List 2011, 272–74); this paper builds on his earlier work with Phillip Pettit: (List and Pettit 2006).

<sup>326</sup> For routine elections in representative democracies, citizens vote on explicit propositions contained in the manifestos or campaign pledges of different candidates. In referenda, they vote directly on the propositions themselves.

a coherent voice even when their opinion is elicited in this highly standardised and regulated manner. List’s contribution focuses on decision procedures with respect to propositions that take the logical form:

$P \text{ iff } Q$

$P$

$\therefore Q$

He uses the following illustrative example:<sup>327</sup>

1. We should invade country X if and only if it has weapons of mass destruction.
2. Country X has weapons of mass destruction.
3. We should invade country X.

He shows that the aggregation of three individually consistent answers to these propositions leads to a collectively irrational result whereby the first proposition is affirmed, the second denied but the third still affirmed.<sup>328</sup>

However, even if binary decision procedures with respect to pre-selected propositional statements were *not* prone to such aggregation problems, we would still have reasons to doubt the extent to which voting procedures effectively channel the views of voters to the government. With regard to a particular decision, one’s first preference may not be on the table. Classical liberals typically feel this way when choosing between the expressed policies of a socially conservative but economically liberal right-wing party and their mirror image in the left-wing alternative. There might also be issues that voters feel are important but over which they cannot express an opinion because they are not included in the manifestos of political parties or, in the case of a referendum, are not listed as an option. Finally, as we saw

<sup>327</sup> (List 2011, 275–76)

<sup>328</sup> The following table, reproduced from List (2011, 276), fleshes out the problem in detail. Persons 1-3 individually provide rationally consistent answers to the above three propositions. However, when their individual answers to each proposition are aggregated using a majoritarian function, we arrive at a set of contradictory collective judgements i.e. we should invade *iff* there are WMDs, there are no WMDs and that we should invade nonetheless.

	“We should invade <i>iff</i> if there are WMD”	“There are WMD”	“We should invade”
Person 1	YES	YES	YES
Person 2	YES	NO	NO
Person 3	NO	NO	YES
Collective Judgements	YES	NO	YES

in chapter one, voting is a poor measure of the *strength* of opinion that voters have over a particular issue.

From a Hayekian perspective, the problem with such democratic procedures is that they exclude “the large body of social knowledge which cannot be put into words”.<sup>329</sup> We already saw Hayek’s emphasis on such knowledge in chapter one. He was joined by many other twentieth century thinkers such as Wittgenstein, Michael Oakeshott and Michael Polanyi in arguing that tacit knowledge is foundational with respect to knowledge expressible in propositional form. For Wittgenstein, understanding the dictionary definitions of words is insufficient and often not even necessary to be proficient in a language. Rather one needs to understand the largely tacit “agreements-in-judgement” that make the use of a word appropriate in one context but not another.<sup>330</sup> For Oakeshott, our explicit and formal knowledge is but a partial representation of a prior ensemble of tacit know-how revealed in action itself. The distinction can be illustrated by the difference between knowing the content of a cook book and knowing how to cook well.<sup>331</sup> Similarly, and using the example of scientific practice, Polanyi argued that expertise in any advanced activity requires a form of connoisseurship irreducible to a formal method. Such knowledge is best gained by doing the activity itself as an apprentice under the tutelage of a master.<sup>332</sup>

The substantive knowledge held by most is tacitly embedded in our practices, routines and forms of life. This extends both to our understandings of value and our understandings of justice. For many if not most of us our conception of the good life is something that we realise through action. It is not the implementation of a theoretically fleshed-out “rational life plan”.<sup>333</sup> Very few of us, philosophers included, conduct our moral life on the basis of a systematic moral theory.<sup>334</sup> Our beliefs as to how we should treat one another are exhibited as much in our conduct itself as in what we can verbally express. These tacit understandings of the good and the right cannot in principle be communicated to government officials via democratic procedures, limited as the latter are to issuing binary judgements on explicit propositional statements. We might imagine that this epistemic opacity would be alleviated

---

<sup>329</sup> (Pennington 2003, 731)

<sup>330</sup> (Wittgenstein 1958)

<sup>331</sup> (Oakeshott 1991)

<sup>332</sup> (Polanyi 2005)

<sup>333</sup> As Rawls assumes when he describes his understanding of a “conception of the good” in chapter seven of *Theory*.

<sup>334</sup> This is probably a good thing since every monist moral theory prescribes very unappealing courses of action in at least some cases. “Morality” is better seen as a complex territory of practices irreducible to a neat theory. On this see Schmidtz (2006).

in small polities or communities in which political leaders and the governed shared the same social practices and circumstance such that the tacit components of their knowledge would be relevantly shared. These conditions do not hold in large-scale social settings.

## 7. The Inevitability of Political Discretion

Even if one believes voters can adequately express their preferences by choosing between binary positions on propositional statements, such statements do not determine the *actions* of government officials. They are indeterminate as to their implementation, which is left to official discretion. One prominent and illustrative example is the 2016 referendum on the United Kingdom's continuing membership of the European Union. The referendum result issued to the UK parliament the following deceptively simple instruction – "Leave the European Union".<sup>335</sup> This instruction has proven, however, to be notoriously indeterminate. Possible interpretations extend from sundering connections with the European Union almost entirely – a "No Deal" – to a host of other conceivable arrangements whereby the UK retains some parts of the existing relationship with the EU but ceases others. These more partial forms of exit admit of many possible permutations depending on the specifics of which arrangements continue and which cease. In other words, the substance of what it means for the UK to leave the European Union has been left to the UK government and Parliament.

An example more familiar within political philosophy might prove useful here too. Suppose a political party wins an election on the pledge to implement Rawls' difference principle or otherwise a policy along the same lines. Let us grant for the sake of argument that this means that the realisation of the difference principle is the "will of the people" which democratic representatives are discharged with carrying out. The difference principle states that social and economic inequalities are required and permitted only to the extent that they maximise benefits to the least-advantaged. Can we conclude that the government have clear instruction as to which policies to carry out? It is true that the difference principle is determinate in some respects. Patently exploitative economic systems, such as feudalism or apartheid, are ruled out. As are policies aiming for, say, a pareto improvement rather than maximising the prospects of the least-advantaged. Still, a great number of institutional possibilities and government policies to achieve them remain ruled *in* given the open-ended way in which the principle might be interpreted. Rawls and his followers believe that a

---

<sup>335</sup> That was the literal wording of the "Leave" option on the ballot paper. The other option was: "Remain a member of the European Union".

radically egalitarian economic system such as “liberal socialism” or “property-owning democracy” best upholds the difference principle.<sup>336</sup> Jason Brennan, however, demurs and argues that a free-market regime would best do so. Part of the disagreement here is simply due to different understandings of the time-frame over which the principle is to be realised. Brennan concedes that over the short-run, greater redistribution might maximise the position of the least-advantaged. However, he argues that a free market regime with comparatively less redistribution would maximise their position over the long-run. The difference stems from the higher long-run growth that Brennan believes is afforded by lower redistribution.<sup>337</sup> Another source of the disagreement, of course, is that Brennan and orthodox Rawlsians espouse different economic theories.

The point here is that the difference principle by itself doesn't specify the economic theory most conducive to its implementation nor what might be the relevant time frame over which it should be realised. Rawls himself recognises this indeterminacy, writing that the judgement as to how the difference principle will be met "frequently depends upon speculative political and economic doctrines and upon social theory generally".<sup>338</sup> He leaves the task of implementing the difference principle to legislators. If the principle can be concretely interpreted in very many ways, potentially up to the point of advocating radically different economic systems, then our elected government have not received any clear instruction as to which policies to carry out. A great deal is left to their discretion. We can generalise this point. Recall from the prior chapter our combining Hayek's theory of local knowledge with Schmitt's emphasis on the unavoidability of a “decisionistic and personalistic” element in the implementation of law. It was argued that the “centre” of the state has limited control over the “periphery”. The substantive content of policy issued by the centre is determined by practitioners at the periphery using their *in-situ* knowledge. That one interpretation of the central policy rather than another is selected is a function of the personality, characteristics, judgements, circumstances and so forth of the particular official-in-question. We may think of these as the political equivalent of the particularistic time and place facts that Hayek identified as pertinent for economic analysis. Such facts are relevant for the phenomena we are trying to understand but which are too varied and idiosyncratic to be accounted for synoptically. We are now applying this notion to the relationship between state officials and citizens. Even if the citizenry unanimously vote for a policy or

---

<sup>336</sup> E.g. (Rawls 2001, 135–38)

<sup>337</sup> (Brennan 2007, 292–93)

<sup>338</sup> (Rawls 1999, 174)

series of policies, their concrete implementation will necessarily require state officials to exercise discretionary judgement on the basis of their “private” if not local knowledge not necessarily available to the demos as a whole.

## 8. The Difficulty of Ex-Post Judgements

One might raise a counter-objection. Even if the implementation of policy requires large degrees of discretionary private authority on the part of politicians, perhaps voters could judge their performance *retrospectively*. The problem with this objection is that it grossly exaggerates the knowledge of individual voters and their capacity to accurately judge the performance of political actors. There is now an established literature on the prevalence of rational voter ignorance. These arguments typically focus on the incentives and motivation – or lack thereof – that voters have to acquire relevant and accurate information when deciding how to vote.<sup>339</sup> The possibility of casting the decisive vote in an election is negligible while the costs of becoming informed are all too real. As a result “most citizens don’t invest effort to be rational about politics because rationality doesn’t pay”.<sup>340</sup> It is not simply that voters are often ignorant of elementary political facts.<sup>341</sup> Even worse, they are often “rationally irrational”. Since nothing is at stake with one’s personal vote, it might pay to engage in politics like a sports fan. Such a fan would rout for their own party or “team” and deride the opposition often regardless of the merits or otherwise of specific policies but because one gains intrinsic satisfaction from the tribal partisanship.<sup>342</sup> The sentimental benefits of such “fandom” outweigh the cost of judging policy matters in an irrational way given the negligible impact one’s vote has on the overall result. If voters, then, tend not only to be rationally *ignorant* but also rationally *irrational* about politics, they are very poorly placed to retroactively judge elected governments, particularly once the scale and complexity of contemporary state activity is taken into account.

We can also identify epistemic constraints rendering ex-post evaluation highly difficult even if voters were highly motivated.<sup>343</sup> I think we can identify three such constraints.

---

<sup>339</sup> Articulations of this argument can be found in Brennan (2016), Caplan (2008), Somin (2013) and Pennington (2011, 65–60).

<sup>340</sup> (Brennan 2016, 49)

<sup>341</sup> For a review of the evidence in the US context, see Somin (2013, chap. 1).

<sup>342</sup> (Brennan 2016, chap. 2); (Caplan 2008); (Pennington 2011, 65–69); (Somin 2013, chap. 3)

<sup>343</sup> I write “highly difficult” rather than “infeasible” so as not to deny the earlier stated point that voters can hold governments to account to some limited extent, namely, with respect to clear and easy-to-recognise policy failures. Thus, one would not be claiming that ex-post voter evaluation of



a) Time scarcity

Chapter two emphasised the diversity of the activities undertaken by state institutions. Recall that even our partial list of these activities ranged from war-making to promoting cultural integration to the administration of the welfare state. Specialist knowledge is a necessary condition of properly evaluating state activity in each of these domains and the many others in which the contemporary state is active. One needs knowledge of economics, public finance and the comparative efficiency of different energy sources to properly evaluate the state's energy policy. One needs to understand different pedagogical techniques to properly evaluate educational policy. One needs to understand international relations, international law as well as military strategy to properly evaluate a contemporary state's decision to go to war. And so on and so forth. Gaining such specialist knowledge is not straightforward. Academics invest a career's worth of time and effort to gain such knowledge. Even their specialisms are normally restricted to one field. The most well-motivated of voters would lack time to gain the full range of specialist knowledge necessary to assess the full range of state activity. Even if they somehow had the knowledge to evaluate *each* and *any* aspect of state activity, they would still lack the time to evaluate *all* or even *most* aspects of state activity since they would have to choose which areas to monitor. We noted in the previous chapter that the UK state contains over 500 distinct bodies and departments. No single voter has sufficient time to monitor the activities of all these bodies and departments even if somehow they had adequate knowledge to monitor any of them in particular.

b) Opportunity cost

Mises and Hayek's seminal contribution to the central planning debate hinged on the claim that the central planner would be ignorant of the opportunity costs of different production plans. This meant that they could not rationally decide between competing uses of the resources under their control. A parallel problem faces voters in assessing the performance of governments. A thought experiment recently advanced by Samuel Decanio is helpful in articulating this point.<sup>344</sup> Imagine a scenario whereby voters have to assess the proposals of different parties for the production of a fuel efficient car. We can even posit, for the sake of

---

governments is impossible in, say, the categorical sense that Hayek and Mises argued was the case with respect to economic central planning. Still, one would be claiming that voter evaluation beyond these easy-to-recognise cases is very difficult due to epistemic constraints that hold regardless of voter motivation.

<sup>344</sup> (DeCanio 2014); though I adapt DeCanio's original proposal for my purposes here.

argument, that these voters have the engineering expertise required to assess whether the elected party actually succeeds in producing the car promised in their proposal. Suppose a party is elected, produces the car they promised and our exceedingly knowledgeable voters are aware of this. A problem still remains. They would not know whether the car was produced efficiently given the *opportunity cost* of resources used in its manufacture.<sup>345</sup> The funds used to produce the car could have been put to countless other uses. They could, for instance, have been used to improve public healthcare or national defence or the education system or indeed been given back to the public in the form of a tax cut. Since the alternative uses for these resources are practically countless, even the most well-motivated and knowledgeable of voters would not be able to assess whether the car was produced efficiently or not.<sup>346</sup>

c) Radical ignorance

Following Sandy Ikeda and Jeffrey Friedman, we might distinguish *radical* from *rational ignorance*. I am rationally ignorant when I know how to acquire knowledge that would improve my state of affairs but do not do so because the expected cost of acquiring it exceeds the expected benefit.<sup>347</sup> Such ignorance is *chosen* and reducible to incentives and motivation.<sup>348</sup> Radical ignorance is not voluntary in this sense. I am radically ignorant if I am unaware of the existence of relevant knowledge that could help improve my state of affairs.<sup>349</sup> Alternatively, I could be aware of the existence of such relevant knowledge but be unaware of how to obtain it.<sup>350</sup> Another form of radical ignorance is *interpretive*. Facts do not speak for themselves but have to be interpreted via theories. However, the right theory with which to interpret complex social phenomena is rarely obvious. The experts themselves typically disagree on such matters. They are subject to Rawlsian burdens of judgement.

When assessing government attempts at governing complex phenomena, even the most conscientious and well-informed voters might not know the most appropriate theoretical

---

<sup>345</sup> (DeCanio 2014, 641)

<sup>346</sup> What is the difference between this case and one where the actor spends their own money? There the countless alternative uses are not a problem because the actor just needs to determine which use of the resource they *personally* think is best. This is insufficient in the context of social choice on behalf of all citizens. There each citizen has to consider alternative uses that could be made of the resources-in-question by a multitude of *other agents* of whose concrete preferences they are unaware.

<sup>347</sup> (Ikeda 2003, 67)

<sup>348</sup> (Friedman 2006, 29- 33); C.f. the aphorism attributed to Einstein: “You don’t have to know everything. You just have to know where to find it.”

<sup>349</sup> (Ikeda 2003, 67)

<sup>350</sup> A PhD student might have good reason to think that an out-of-print book would help with their project but be unaware of how to find it.

framework with which to base their evaluation. Take the example of diagnosing the cause of the 2008 financial crisis. Some experts think it was caused by implicit guarantees of government bail-outs.<sup>351</sup> Some locate the problem in lax monetary policy.<sup>352</sup> Some find it in legal privileges protecting large financial rating agencies from competition.<sup>353</sup> Others in the 1999 repeal of the 1932 United States Glass-Steagall Act, after which American banks no longer had to separate their commercial and investment banking divisions.<sup>354</sup> Still others blame “market-to-market” accounting standards.<sup>355</sup> The difficulty even experts face in obtaining a theoretical handle on phenomena as complex as the financial crisis is not surprising from a Hayekian perspective. The sheer number of actors combined with the interdependent nature of their behaviour makes it very difficult to properly understand the causal relationships between actor behaviour and outcomes. This is a key feature of emergent properties. By the same token, our inability to understand such phenomena inhibit our ability to evaluate the state agents who attempt to regulate them.

So ends this chapter’s critique of the principal-agent model. To recap, we have argued against it on three grounds. First, tacit voter knowledge and values cannot be transmitted to the government. The transmission of public opinion is limited to a binary assessment of propositional statements. Second, even if we posit that the electorate affirm a policy as expressed in propositional terms, its implementation is subject to the discretion of state officials on the basis of their private knowledge. Third, various forms of radical voter ignorance and also incentive problems mean that ex-post evaluation on the part of voters cannot alleviate the problem of discretionary authority. We conclude that state actors cannot feasibly be understood as an agent of the citizenry.

## **9. On the Value of Democracy**

This chapter and this thesis is critical of the idea of collective political agency, a concept naturally linked to democracy. It thus behoves this thesis to clarify the relation of its argument to the value of democracy. Such is the task of the final section of this chapter. While the arguments advanced in this work are in tension with claims as to democracy’s intrinsic value, they are quite compatible with claims as to its instrumental value and even

---

<sup>351</sup> (Roberts 2019)

<sup>352</sup> (Taylor 2009)

<sup>353</sup> (Friedman 2006)

<sup>354</sup> Including, perhaps surprisingly, the libertarian Loren Lomasky (2011).

<sup>355</sup> (Kothari and Lester 2011)

its epistemic value. One might think that democracy has intrinsic value because democratic participation constitutes the exercise of a powerful form of collective autonomy. Quite obviously, the arguments of this chapter and thesis run against this justification. Another distinct reason why democracy might have intrinsic value is that democratic procedures show equal respect to all citizens under circumstances of deep disagreement about both justice and the good. Given such disagreement, it is argued, granting each party an equal say in the determination of laws is required as a matter of procedural justice.<sup>356</sup> This is distinct from the collective autonomy view because to say I was afforded equal consideration in a process is not necessarily to say that the outcome of that process would accord with my will. The arguments of this thesis and this chapter are also in tension with this understanding of democracy's value, which presumes the validity of the associational view of political society. We argued in chapter one that liberal citizens are too ignorant of one another to be characterised as in deep and widespread disagreement. Furthermore, the idea of equal determination in the formulation of laws is a chimera if this chapter's argument as to the inevitability of discretionary governmental authority is correct.

Instrumental arguments for democracy locate its value in the outcomes it produces.<sup>357</sup> On the instrumental view, democracy is desirable if it is better at producing these outcomes than alternative regimes. One powerful case for the instrumental value of democracy rests on its capacity to prevent domination. Samuel Bagg puts this point well: "[though] hardly realizing the classical dream of collective self-rule, elections provide some leverage over leaders, as well as non-violent means for overthrowing would-be tyrants".<sup>358</sup> The point here isn't simply that democracy gives voters the capacity to, as it were, throw the rascals out.<sup>359</sup> The right to vote also provides some measure of protection against the state's worst excesses, particularly for minorities or otherwise marginalised groups.<sup>360</sup> Democratic institutions' comparative advantage in avoiding calamities such as famines should, I think, be understood in this light. I do not see any incompatibility with the argument of this thesis and such instrumental justifications.

---

<sup>356</sup> Christano (2008), Waldron (2004) and Valentini (2013) defend versions of this position.

<sup>357</sup> Arneson (2004) and Brennan (2016, 10–14) defend this view, although Brennan believes that a limited suffrage restricted to more knowledgeable voters, an "epistocracy," might prove instrumentally superior to democracy.

<sup>358</sup> (Bagg 2018, 901–2) c.f. "Democracy is essentially a means, a utilitarian device for safeguarding internal peace and individual freedom." (Hayek 2001 [1944], 73)

<sup>359</sup> This, for Hayek, was the chief value of democracy. See Kukathas (1998, 24).

<sup>360</sup> Jacob Levy (2017, 116–18) makes this very point in a paper otherwise strenuously criticising the idea of democratic unity.

Epistemic defences of democracy, meanwhile, present a case for it akin to the Hayekian case for the market economy. Drawing on the Hong and Page Theorem that diverse groups are better at problem solving than comparatively more intelligent but homogenous groups, H  l  ne Landemore argues that democratic decision-making can produce wise results even if individual voters are not well-informed about politics.<sup>361</sup> Her model holds only under special conditions and it is doubtful that they are met in real-world elections.<sup>362</sup> Still, supposing that the conditions were met, this would not be inconsistent with the arguments of this chapter and thesis. Democratic outcomes do not reflect any collective intention on Landemore’s theory. Rather they reveal independent truths by aggregating the discrete opinions of many voters. They are still emergent in the relevant sense of not being straightforwardly reducible to the individual opinions of the voters-in-question.<sup>363</sup>

## Conclusion

Our critique of descriptive associationalism is now complete. The arguments presented here do not constitute a conclusive case against the associational account. They merely show that the associational model does not accurately describe actually-existing liberal democracies. Even if the associationalist accepted this, the possibility remains that collective political agency would be possible under more propitious circumstances. In the next chapter we will consider whether associationalism might be possible if we assumed that all citizens were committed to a Rawlsian understanding of justice.

---

<sup>361</sup> (Landemore 2017)

<sup>362</sup> As observed by Brennan (2014) and Bagg (2018, 894).

<sup>363</sup> If we already knew *a priori* the truths revealed by democratic procedures, then democracy would lose its *raison d’  tre* if justified solely on epistemic grounds. The truth “revealed” by democratic procedures have to be emergent then in this sense also.



## Chapter Four: Aspirational Associationalism – The Well-Ordered Society

### Abstract

The last chapter argued that mutual ignorance prevents citizens of actually-existing liberal democratic societies from carrying out joint-intentional action. This chapter argues that the problem would remain even if we assumed all liberal citizens individually affirmed Rawlsian justice. For they would still not be able to reliably *signal* this commitment to one another under impersonal conditions. This crucial condition necessary for joint-intentional action would still not be met. This is the case whether the signalling mechanism is deliberation or voting. The basic difficulty is that one needs reliable knowledge of the internal motivation of another agent to know if they are truly committed to liberal justice. We can gain that knowledge only of a small subset of our peers in an impersonal liberal society.

### Introduction

The Rawlsian would likely be unmoved by the arguments of the prior chapter. Even if that argument succeeds, they might argue it does little to budge Rawls' own view. The fact that collective political agency is infeasible under actually-existing liberal democratic societies does not show it to be infeasible under the idealised conditions of Rawls' well-ordered society. This chapter then considers whether such a society is possible. Let us recall the two key desiderata to the well-ordered society:

1. Everyone accepts and knows that the others accept the same principles of justice.
2. The basic social institutions generally satisfy and are generally known to satisfy these principles.<sup>364</sup>

We have already mustered arguments against the feasibility of the second desideratum. Chapter two argued that there is no agency who can redesign the social institutions according to a fixed blueprint. The authority to create and revise major institutional rules in liberal democratic societies is de facto polycentric. Just as no agency is suitably positioned to centrally plan economic production and distribution, none are suitably positioned to centrally plan the major institutional rules of society. If the arguments of chapter two are sound, there is no clear route from the status quo to a well-ordered society as conceptualised by Rawls. Our focus in this chapter, however, is on the first desideratum. The stability and maintenance of the well-ordered society is a function of the shared intentions of the liberal citizenry. When thinking about this issue, Rawls implicitly understood the claim of Bratman

---

<sup>364</sup> (Rawls 1999, 5)

and others that mutual assurance is a necessary condition of joint action. Rosie and Jim share an intention that they *X* iff each knows that the *other* intends that they *X*. *Mutatis mutandis*, liberal citizens intend that they give justice to one another iff each knows that all others intend that they give justice to one another. Rawls' understanding of this mutual knowledge requirement expresses itself in the first desideratum of the well-ordered society. Each citizen must know that every other citizen affirms the principles of justice. If the ideal of a well-ordered society governed by the shared will of liberal citizens is to be met, liberal citizens need to obtain assurance of one another's intentions. This is Rawls' assurance problem.

This chapter will argue that this mutual assurance problem cannot be solved in a large-scale liberal society. The arguments and thought-experiments of this chapter are consistent with the possibility that liberal citizens are all *individually* committed to Rawlsian justice. What they demonstrate is the difficulty facing any particular citizen in trying to assure themselves of this. The problem is that it is very hard to gain credible assurance of the intentions of the mass of our peers in an impersonal society. Even if we grant to Rawls the idealising assumption that all citizens individually affirm Rawlsian principles, then, their inability to credibly communicate this fact to one another prevents them from coordinating their actions on the basis of his conception of justice. We thus demonstrate that the epistemic constraints on collective action remain even if we idealise away the motivational ones.

The argument to follow is divided into five sections. The first details Rawls' assurance problem. It lays great stress on the *legislative* nature of this assurance problem. Since citizen compliance with laws is coercively enforced, it is only with respect to their freely-made decisions to vote that they can gauge their respective intentions to maintain just institutions. This fact has not been sufficiently appreciated in the secondary literature. There are two possible signalling mechanisms whereby citizens can make clear to one another whether they intend to vote justly. The first is *deliberation*. Citizens try to assure one another of their commitment to voting on just terms by using only the discourse of justice in political deliberation. The second is the act of *voting* itself. The fact of voting for a just party might itself be indicative of one's commitment to liberal justice. Sections two and three consider both these mechanisms and find them wanting. Sections four and five consider and rebut respectively two objections to the chapter's argument.

## **1. The Rawlsian Assurance Problem as a *Legislative* Assurance Problem**



Rawls assumes individuals have two powers.<sup>365</sup> Their “rational” power is to advance their personal but partisan conception of the good. Their other power is a sense of justice such that they are willing to comply with fair terms of cooperation with their peers. In the well-ordered society, all liberal citizens have a sense of justice strong enough not to be overridden by their partisan understandings of the good. Suppose I am a dutiful citizen of such a society. Faced with a choice between a theocratic party that will impose on all the demands of what I consider to be the true religion and a Rawlsian party that will secure religious freedom for all, my sense of justice will direct me to vote for the latter. Citizen motivation to comply with liberal justice, however, is contingent upon reciprocation. For when “any one person knows or reasonably supposes that others will not do their part, it will be worth his while to be the first, or among the first, not to do his, or even dangerous for him not to be”.<sup>366</sup> Our participation is contingent upon reciprocation because we desire not to be taken advantage of. I will dutifully pay my taxes if I know everyone else will but not if I am the only “sucker” who pays while others free-ride; I will dutifully endorse legal toleration of views I find abhorrent but not if adherents of those views refuse to endorse legal toleration of my views; and so on. Just as Bratman’s painters have to assure one another that they will show up on the day and are not shirking during the painting process, Rawlsian citizens have to assure one another that that they won’t defect from doing their part in fulfilling the common objectives of their political association. It is now common to model Rawls’ reciprocity requirement as an assurance game from game theory.<sup>367</sup>

---

<sup>365</sup> E.g. (Rawls 1999, 18–25)

<sup>366</sup> (Rawls 1963, 290)

<sup>367</sup> This interpretation was pioneered by Paul Weithman (2011, chap. 2) c.f. (Kogelmann and Stich 2016, 717–20); (Thrasher and Vallier 2013); (Gaus 2011a)

		Jim	
		Comply	Defect
Rosie	Comply	1 <sup>st</sup> / 1 <sup>st</sup>	2 <sup>nd</sup> / 4 <sup>th</sup>
	Defect	4 <sup>th</sup> / 2 <sup>nd</sup>	3 <sup>rd</sup> / 3 <sup>rd</sup>

**Figure 4.1 – Rawlsian Mutual Assurance Game**

Rosie and Jim are Rawlsian citizens with a well-established sense of justice. Their pay-offs are numbered ordinally with the 1<sup>st</sup> being the highest and the 4<sup>th</sup> the lowest. Each gains the highest pay-off when both comply.<sup>368</sup> Their lowest pay-off arises when they comply but the counter-party defects i.e. they are the “sucker” who is taken advantage of. Unlike in the prisoner’s dilemma, each party does not have a dominant strategy regardless of what the counterparty does.<sup>369</sup> If the counterparty complies, their best strategy is to do likewise. If they defect, their best strategy is also to do likewise. There are, then, two Nash equilibria: the risk dominant one whereby both defect and the pay-off dominant one whereby both comply. The challenge for liberal citizens is to assure one another of their commitment to liberal justice so as to reach the pay-off dominant equilibrium. The assurance problem of the well-ordered society might be considered as a (very high) N-person version of this simple model.

At this juncture, however, we face a complication. For on my reading Rawls posits two different kinds of assurance games for liberal citizens to solve and only one of them stands a chance of supplying the information required to meet the first desideratum of the well-ordered society. There are two distinct ways to be just in a Rawlsian society. One can comply with just laws and one can vote for just laws. The *compliance assurance problem* arises because citizens render their compliance with just laws contingent on reciprocation. I will dutifully pay my taxes but only if I know that my peers will do likewise. Rawls’ solution to this assurance problem is Hobbesian. Citizens in a large-scale community have no effective

<sup>368</sup> Note that ordinal utility rankings in this context simply mean that the agent-in-question ranks one state of affairs higher than another. A utility function need not presume that agents are egoistic or utilitarian, as is often assumed. On this see Gaus (2012a, 158–60).

<sup>369</sup> In the prisoner’s dilemma, the dominant strategy of both players would be to defect. The prisoner’s dilemma is displayed in figure 7.1 of chapter seven.

means of signalling to their peers their willingness to pay tax, since “the degree of mutual confidence in one another’s integrity that renders enforcement superfluous is not be expected”.<sup>370</sup> However, the “mere existence of an effective sovereign” believed to be capable of detecting and punishing defectors is enough to create confidence that one is not going to be a sucker when they pay their taxes.<sup>371</sup> Note, however, that if Rosie’s compliance with justice is coercively enforced, it is not clear whether Rosie complies to be just or simply to avoid punishment. Thus, this assurance mechanism fails to transmit the information necessary to satisfy the first desideratum.

The *legislative assurance problem* arises because citizens render their willingness to vote justly contingent on reciprocation. This is exemplified in the voter who would vote for liberal toleration over a theocratic party that favoured his religion so long as other voters would be similarly restrained in their voting behaviour. On my reading, this is the primary assurance problem with which Rawls is concerned. There is exegetical evidence for this claim. In a significant passage in *Political Liberalism*, Rawls refers to the importance of liberal citizens “acting willingly so as to give one another justice over time”. It is in this passage where Rawls explicitly argues that stable liberal institutions should result from the willed actions of liberal citizens rather than dictatorial imposition. Crucially, he writes that this kind of stability “connects” with the idea that sovereign political power in a liberal society “is the power of equal citizens as a collective body”.<sup>372</sup> In his 1997 paper, “The Idea of Public Reason Revisited,” he says outright that the “ideal of public reason” is realised by liberal citizens in their voting conduct.<sup>373</sup> This textual evidence suggests that for Rawls citizens show their willingness to give justice to one another primarily in their capacity as legislators of the law rather than in their capacity as subjects obligated to comply with the law.

However, the main argument for this claim is conceptual. Only this assurance mechanism is a viable candidate for meeting the first desideratum of the well-ordered society. If the latter is to be met, citizens need credible knowledge of one another’s sincere commitment to liberal justice and such knowledge can come only from freely exercised actions rather than coerced ones. The right to vote is by definition freely exercised. Jim’s best chance, then, of inferring whether Rosie really shares his commitment to realising or maintaining a just society is by observing her uncoerced voting activity. If he could satisfy himself that she did

---

<sup>370</sup> (Rawls 1999, 237)

<sup>371</sup> (Rawls 1999, 237)

<sup>372</sup> (Rawls 1996, 143)

<sup>373</sup> (Rawls 1997, 769)

vote justly, then we have a mechanism that could meet the first desideratum of the well-ordered society. The trouble, as we shall see, is that there is no straightforward way of inferring whether someone has defected from liberal justice by observing which candidate or what policy they voted for. One has to know *why* they voted for that person or policy. However, in a large-scale society, we can only gain knowledge of the motives of a small subset of our peers. A couple clarificatory points are in order before we commence with the substantive argument.

What is the relevant content of liberal justice that citizens are united in realising? Rawls gave two answers to this at different points in his career. The Rawls of *Theory* and the Rawls of the initial phase of his political turn assumed that all members of the well-ordered society would share the intention to realise justice as fairness. However, in the later, paperback edition of *Political Liberalism* and subsequent papers, Rawls begins to think this too much to hope for. Reasonable disagreement would extend to questions of justice as well as the good. The later works exhibits a thinner understanding of the shared intentions binding liberal citizens. The well-ordered society can no longer be one where citizens are united by the joint-intention to realise justice as fairness. They are still united by the joint-intention to manage their political association on the basis of a “reasonable” conception of justice though they can now legitimately disagree as to which is the *best* conception. A reasonable conception of justice has to contain the following three features:

First, a list of certain basic rights, liberties, and opportunities (such as those familiar from constitutional regimes);

Second, an assignment of special priority to those rights, liberties, and opportunities, especially with respect to the claims of the general good and perfectionist values.

Third, measures ensuring for all citizens adequate all-purpose means to make effective use of their freedoms.<sup>374</sup>

The key idea is that a reasonable conception of justice constitutes reasons at least in principle justifiable to all. I do not invoke a partisan reason when I invoke such a conception in political discourse, in contrast to when I invoke, say, my personal religious values. Still, according to the later Rawls, we disagree on which conception of justice is most reasonable in this regard. Justice-as-fairness is but one of a family of conceptions of justice that citizens might affirm.<sup>375</sup> A law is just if its rationale is consistent with any of these reasonable conceptions. Such a

---

<sup>374</sup> (Rawls 1997, 774)

<sup>375</sup> (Rawls 1997, 774 fn27)

rationale is what Rawls calls a “public reason”. Citizens, then, are just when they vote on the basis of public reasons but there is no expectation that all will vote on the basis of public reasons derived from the same conception of justice. I will, for the most part, assume this view of liberal justice in the analysis to follow since it is Rawls’ considered view. Putting all this together, the question remaining before us is whether impersonally related liberal citizens can assure one another of their shared intentions to govern their shared political association the basis of public reason.

We should note an exegetical difference between my reading of Rawls and that of Kogelmann and Stich, who also interpret Rawls as conceiving of two assurance problems.<sup>376</sup> Though this chapter’s account of the two assurance problems differ from theirs. On their interpretation, there is a *citizen assurance problem* analogous to the above compliance assurance problem. This holds that each citizen will comply with just laws so long as penal institutions exist to assure them that their peers will do the same. However, Kogelmann and Stich also claim that normal citizens have no obligation to comply with the norms of public reason other than obeying just laws. On their reading, only elected politicians and judicial officials need do so. The norms of public reason solve a *public official assurance problem*. Elected politicians and judicial officials have to assure one another that they will legislate or issue legal decrees in accordance with public reason. They send this assurance by using public reason in legislative debates, justifying legal rulings with public reason, and so forth. Kogelmann and Stich’s account of Rawls’ view may be summarised as one whereby politicians and judges deliberate upon and produce just laws while citizens obey those laws. This assumption allows them to say some interesting things about the conditions under which assurance problems might be solved. However, this is unlikely to be Rawls’ position. For if it is only public officials who show their commitment to public reason, while the rest of us are coerced into it, again no assurance is gained that the citizenry as a whole have a commitment to liberal justice. As such, if Kogelmann and Stich’s reading is correct, it is left unclear how Rawls intended the first desideratum of the well-ordered society to be met. Furthermore, already cited earlier was Rawls’ own explicit statement that the norm of public reason does apply in the voting booth.<sup>377</sup>

---

<sup>376</sup> (Kogelmann and Stich 2016, 719)

<sup>377</sup> (Rawls 1997, 769)

## 2. Deliberation as a Signalling Mechanism

The first signalling mechanism is a deliberative one whereby citizens use only public reasons when advocating public policies. If Jim presents arguments in favour of a public policy using public reasons then his fellow citizens could conceivably interpret this as a signal that he intends to vote on such a basis. The existing literature has focused on this as the signalling mechanism, and many have already argued that it is ineffective.<sup>378</sup> One difficulty here is that even if it were an effective mechanism, it is hard to see how it would solve the assurance problem unless public deliberation was a very popular activity among liberal citizens. This, as Rawls himself believed, seems an unrealistic expectation in large-scale democracies.<sup>379</sup> Be this as it may, and in keeping with focus of this thesis on epistemic rather than motivational problems, let us grant the widespread use of deliberation to see why it is an ineffective voting mechanism. This section will review and expand upon some of the arguments already developed in the secondary literature. There are three key problems.

### (i) The problem of cheap talk

This criticism is raised by Gerald Gaus.<sup>380</sup> We can state his criticism with extra perspicuity if we bear in mind the legislative nature of the assurance problem, which Gaus himself does not consider. Rosie looks for assurance that Jim intends to vote on the basis of public reasons rather than his own partisan reasons. Jim tries to reassure Rosie by invoking only public reasons during their political deliberations. However, this constitutes a weak signal for the following thought might well strike Rosie. *If* Jim did lack civic virtue, he would happily benefit from others exercising their vote in accordance with the restraints of public reason while himself voting according to his own partisan reasons. Further, he would be incentivised to *merely* deliberate according to public reasons to falsely signal his commitment to it. If the false signal worked, Rosie would be mistakenly assured to vote on the basis of public reasons but Jim could still vote on the basis of his own partisan reasons in the privacy of the voting booth. He is incentivised to send this signal because talk is cheap. Rosie, being aware of this, is thus *not* reassured by Jim's invocation of public reasons during deliberation.

### (ii) The problem of reflexivity

---

<sup>378</sup> (Thrasher and Vallier 2013, 941); (Gaus 2011a, 316); (Kogelmann 2019)

<sup>379</sup> (Rawls 2001, 143)

<sup>380</sup> (Gaus 2011a, 317)

Suppose Rosie does believe Jim’s signal to be credible. Another problem remains. Gaus also observes that signalling mechanisms need also overcome the problem of reflexivity.<sup>381</sup> It is not enough that Rosie accepts Jim’s signal. Jim additionally needs to know that she has received and interpreted the signal appropriately. She, in turn, needs to know that he knows she has received and appropriately interpreted the signal. The problem can go on *ad infinitum*. We saw in the previous chapter that this issue also arises with Bratman’s model. Successful solutions to assurance problems require a great deal of common knowledge between the participants. Familiarity and iterative reactions can overcome this problem at a small-scale.<sup>382</sup> If Rosie and Jim meet regularly, then it will soon become common knowledge between them that they have received one another’s signals. These conditions do not hold in large-scale settings where most of our deliberative partners are strangers.

### (iii) Noise

Thrasher and Vallier identify noise as another problem afflicting deliberation in a large-scale setting – “Noise is the problem of distinguishing between communication by citizens that signal allegiance to the public conception, and hence assurance, and forms of communication that do not”.<sup>383</sup> Thrasher and Vallier stress that that the permissive nature of the public reason makes it very prone to the problem of noise. For example, Rawls says citizens are allowed to invoke non-public reasons in deliberative public reason so long as they meet the following proviso, “that in due course proper political reasons – and not reasons given solely by comprehensive doctrines [i.e. non-public reasons] – are presented that are sufficient to support whatever the comprehensive doctrines introduced are said to support”.<sup>384</sup> Rawls is deliberately vague about the specifics: “the details about how to satisfy this proviso must be worked out in practice and cannot feasibly be governed by a clear family of rules given in advance”.<sup>385</sup> However, if this is allowed, it is very difficult in the “cacophony of the public sphere” to distinguish genuine defections from those speech-acts where the public reasons are eventually forthcoming.<sup>386</sup>

---

<sup>381</sup> (Gaus 2011a, 318)

<sup>382</sup> (Gaus 2011a, 318)

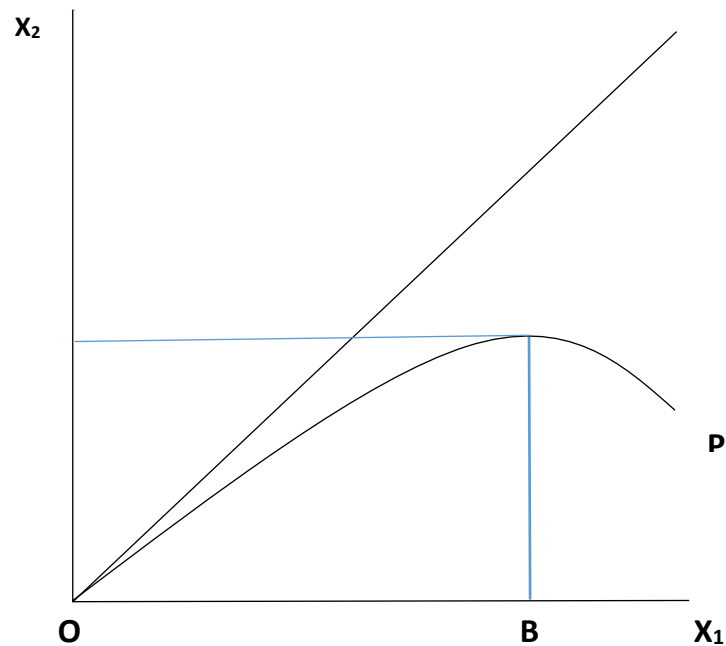
<sup>383</sup> (Thrasher and Vallier 2013, 942)

<sup>384</sup> (Rawls 1997, 784); Rawls was forced to make this concession to avoid the rather embarrassing implication that his theory would otherwise cast the abolitionist and civil rights movements as “unreasonable” for their decidedly theological discourse.

<sup>385</sup> (Rawls 1997, 784)

<sup>386</sup> (Thrasher and Vallier 2013, 942)

One might be tempted to think that Thrasher and Vallier’s noise problem afflicts only the later more permissive understanding of liberal justice. This would be incorrect. For even in *Theory* Rawls allows for disagreement in the interpretation of justice as fairness.<sup>387</sup> This would cause noise problems even on the much more restrictive view of liberal justice found in that early work. Let us take the difference principle again. A diagrammatical representation of it, figure 3, is reproduced from *A Theory of Justice*.<sup>388</sup>



**Figure 4.2 – The Difference Principle Contribution Curve**

The axis  $X_1$  denotes the income of the most advantaged and  $X_2$  correspondingly the least advantaged. As is very well known, the difference principle holds that departures from the initial point of equality  $O$  are justified so long as they make  $X_2$  better off. Hence, in this highly stylised example, the principle would determine that the income of  $X_1$  should be allowed to rise up to point  $B$ , at which the corresponding come of  $X_2$  is maximised. Any further marginal increase in the income of  $X_1$  will lead to corresponding marginal reductions in the income of  $X_2$ . Now, of course, this model is highly simplified. It holds all other factors constant, which is not true of actual economies. All manner of other variables can affect the income of  $X_2$  apart from the income of  $X_1$ . Additionally, in the real world we will not know for certain where we are on the contribution curve  $OP$ . This means that it will not always be obvious

<sup>387</sup> (Rawls 1999, 170)

<sup>388</sup> (Rawls 1999, 67)



which policies will best realise the difference principle. This question is subject to the burdens of judgement.<sup>389</sup>

Suppose now that in a deliberative forum one group of citizens argues for a tax cut. They argue that the economy is currently positioned to the left of B on the contribution curve. Their argument is that tax cuts will incentivise more economic activity on the part of the most advantaged and, through a resulting combination of Laffer curve effects and higher economic growth, move us closer to B thus improving the position of the least advantaged. Suppose also that many of these citizens are themselves in the most advantaged group. It is not clear in such cases how one distinguishes sincere views as to how to further Rawlsian justice from more self-serving claims. These citizens could have a genuine sense of Rawlsian justice and sincerely believe we are to the left of B on the contribution curve. Or they could just want to reduce their personal tax bills. In which case they are taking advantage of our uncertainty as to where we are on the contribution curve to generate a convenient justification. Thus, the problem of noise cannot be solved by merely making the requirements of liberal justice or public reason stricter. For even if public reason is restricted to a unitary conception of justice, ambiguities in its application would create noise. Note again, as with the reflexivity problem, it is not difficult to imagine how the noise problem might be overcome in small groups where the deliberators get to know one another well and there are many recurring opportunities for deliberation. Local knowledge of one another might build trust in such settings. However, such local knowledge is lacking in impersonal settings.

### **3. Voting as a Signalling Mechanism**

If deliberation fails to signal our commitment to public reason, perhaps the act of voting itself can do the trick. By voting for candidates that govern on the basis of public reason or punishing those candidates that fail to do so, I signal to my peers that I participate in our joint-intention to realise a well-ordered liberal society. Unlike deliberation, the established literature has focused little on the act of voting itself as a signalling mechanism. Thus here we stake out a more independent path with our analysis. Unfortunately, the act of voting itself is, if anything, an even worse signalling mechanism than deliberation. It is much harder than it appears to gauge voter intentions from their voting behaviour. Consider the following hypothetical case of an elected politician.

---

<sup>389</sup> As we already saw in the prior chapter.

Yesterday was Election Day. A candidate was running for office on the promise of restricting abortion rights. This campaign pledge was announced a month ago and it was backed by theological rather than public reasons. Today the results have been announced and the candidate has won her seat by a healthy majority. She has not thus far made any effort to satisfy the Rawlsian proviso.<sup>390</sup>

Suppose Rosie voted for our politician and that our good Rawlsian citizen Jim considers this to be clear-and-cut case of defection on Rosie's part. He might think that a month must be more than enough time to meet the proviso. It can't take that much time for our politician to come up with *bona fide* public reasons for that policy. If she hasn't met the proviso in the month between the day of the pledge and Election Day, then it is clear to Jim that she has no intention of doing so. He expects Rosie to have reached the same conclusion about the politician's intentions regarding the proviso. Therefore, thinks Jim, Rosie's vote for the politician signals her defection from public reason.

Jim's conclusion, however, would be too hasty. Note that Rawls himself says that the specific details of the proviso – “when does it need to be satisfied? On the same day or some later day? Also, on whom does the obligation to honor it fall?” – cannot be worked out in advance but need to be left to the “good faith” of the parties.<sup>391</sup> This seems sensible. Further, upon some more reflection it is not obvious that a month is too long to meet the proviso. One might take the view that the proviso is an important but technical matter to be dealt with at a later stage. Perhaps, on this alternative view, the most appropriate time to offer public justification is when policy is turned into explicit legislation rather than during pre-election campaigning. If one could reasonably take this alternative view of the proviso's requirement, then Rosie's voting behaviour does not straightforwardly speak for itself as Jim might have initially assumed. It *could* be the case that she has voted for the politician simply because she has defected from public reason. It could also be the case that Rosie has not defected but takes this alternative view of what the proviso requires and has voted for the politician in good faith that the proviso (as Rosie understands it) will be fulfilled at a later date. The bare fact of her voting decision is indeterminate with respect to these two conclusions.

---

<sup>390</sup> Note that Rawls (1996, 243 fn. 32) says public reason itself doesn't pre-judge the abortion issue one way or another but stipulates that any policy position with respect to it should be justified on the basis of three “reasonable” values: respect for human life, the ordered reproduction of political society and the legal equality of women. Kogelmann (2019, 205–8), in a different paper to his aforementioned one, observes that depending on how these three values are ranked, one could arrive at the full gamut of positions with respect to abortion ranging from complete prohibition to no or very limited restrictions.

<sup>391</sup> (Rawls 1997, 784)

These interpretive difficulties are only compounded the more realism we introduce into our analysis. The complexity of voting systems adds one layer of interpretive difficulty. Under complex voting systems, *strategic* voting may be worthwhile. Citizens voting under such complex systems may knowingly vote for parties whose policies may be deemed unambiguously unjust from a public reason perspective. Still, the intentions of the voters may credibly be to realise Rawlsian justice. Under the UK's first-past-the-post system, for instance, votes count only insofar as they contribute to that party's tally in the constituency in question. The party that gains a majority of the votes in each constituency will represent that constituency in the House of Commons, with the other votes having no other input in the democratic procedure to decide the government. The party that wins a majority of constituencies is entitled to form the national government.

Suppose Rosie and Jim are British citizens. We might also suppose, as one might imagine Rawls would, that the policies of the British Labour Party are most consistent with public reason. Further, that the British Conservative Party is the second most consistent with public reason, and (if I am allowed some uncreative licence) the newly-formed Unreasonable Party the third and least most consistent. Rosie lives in a constituency where Labour have no chance of winning. It is, instead, a two-horse race between the Conservative Party and the Unreasonable Party. The national election, however, is a two-horse race between the Conservative and Labour parties, such that the Unreasonable Party have next-to-no chance of winning enough constituencies to form the next government. Rosie votes for the Unreasonable Party in her constituency. This fact alone, however, doesn't mean she has defected from public reason. For under such conditions, it is rational for a supporter of the Labour Party to vote for the Unreasonable Party to reduce the possibility of the Conservative Party winning that constituency and in turn increasing the chances of the Labour Party forming the next government nationally. So the *input* of voting for the most unjust party locally might reasonably be intended to realise the *output* of realising public reason nationally. As far as Jim is concerned, Rosie has defected from public reason if she voted for the Unreasonable party out of approval for their policies but not if she voted for this strategic reason. Again, he cannot infer either way from the act of voting itself.

Another complication is the fact that parties campaign on and indeed enact bundles of policies rather than singular policies. Some of these may be just and some unjust. The fact of bundling heightens the interpretative problem in two respects. The first stems from non-culpable voter ignorance. Even if we posit Rawlsian citizens with a strong sense of liberal justice, it seems unrealistic to expect them all, for instance, to take graduate degrees in social

science to help them better evaluate government performance. It also seems unrealistic to expect them all to spend endless hours reading campaign manifestos or endless hours monitoring the activities of their legislature. Voters might reasonably make their decisions only on the basis of the party's headline policies or on the basis of issues that they are most knowledgeable about. They might vote, then, on the basis of what economists call imperfect information. Under such conditions, they might well be *reasonably unaware* of the unreasonable policies of the party who they vote for or otherwise the unreasonable implications of some policies. To the extent that this is true, Jim cannot infer from Rosie's voting for a party whose bundle of policies is not completely reasonable the conclusion that her commitment to public reason is flagging. For she may have decided to vote for that party only on the basis of their more reasonable policies of which she is aware.<sup>392</sup>

A second interpretive difficulty arises if voters have to make trade-offs of "reasonability" between parties whose bundles of policies are at most partially just. Imagine two parties. The policies of party A unambiguously increase justice on one dimension – say foreign policy – but also unambiguously increase injustice on another dimension – say penal law. The policies of party B maintain the status quo. Rosie votes for party A. Does Jim infer from this that she has defected from public reason? Perhaps if she has no interest in foreign policy and actively *favours* the unreasonable penal policy. But she could also be a good Rawlsian who simply thinks the gains in foreign policy *outweigh* the losses in penal law. There is also a further complication here, akin to the proviso, of what the Rawlsian might consider to be reasonable trade-offs when these values are at stake. There seems no obvious way to strike these trade-offs. What might seem to Jim a very unreasonable way of doing so might appear differently to Rosie.

Finally, citizens may vote to realise aims other than affecting legislation, such as communicating their sentiments and concerns.<sup>393</sup> This is a popular explanation for the recent

---

<sup>392</sup> I have been, for the sake of argument, assuming that it is relatively easy to establish which policies are in line with public reason and which not. Of course this is not true in real life. I submit that even in the well-ordered society, citizens might reasonably misunderstand the nuances and complexities of what counts as public reason. This concept, like the rest of Rawls' ideas, is complicated. Certainly, the voluminous commentary his work has inspired indicates how, despite the perspicuity with which Rawls tries to summarise his ideas, they are not straightforward to political philosophers, let alone the average voter who has other concerns in life. A voting decision that looks like a sure fire defection to the committed Rawlsian might just be an honest mistake on the part of the ordinary citizen.

<sup>393</sup> An observation made by Gaus (2009, 15). The Rawlsian might be tempted to say it is unreasonable to vote on the basis of motivations other than enacting legislation. But, as Gaus points out, this seems to "truncate the political, conferring sole legitimacy on one of its functions" (2009, 16). Plus, it is arguably the denial of expressive voting as legitimate that seems unreasonable. Given the minimal

electoral successes of populist parties in liberal democracies. Voters are less interested in the concrete policy platforms of these parties, which are often quite vague, and more so in expressing discontent and sending a message to “The Establishment”.<sup>394</sup> This kind of voter motivation creates further interpretive difficulties. Let us grant that the policies of Donald Trump are inconsistent with Public Reason. Suppose now that our good Rawlsian Jim is aghast that so many of his fellow citizens are voting for Trump. Rosie votes for Trump. Can he infer from Rosie’s decision to vote for Trump that she is unreasonable? Again, matters are not nearly as obvious as they might appear. It could be that Rosie has voted for Trump because she wants his policies realised. In that case she would have defected. It could also be that Rosie voted for Trump to try and communicate her dissatisfaction with the economic policies of the Federal Government. It is not difficult to imagine messages that Rosie might wish to communicate which are consistent with public reason and pleasing to Rawlsians.<sup>395</sup> For example, she might vote for Trump to communicate her belief that the major social institutions of the US are rigged to benefit the most advantaged at the expense of the least advantaged, and that the political class are complicit in this exploitative process. Given the complicity of the established parties in this process, one way Rosie might express her (by hypothesis) reasonable concerns is by voting for an outsider. If this was her primary if not sole reason for voting for Trump, then it is far less clear that she has defected from public reason. If her aim is not so much to have Trump’s policy’s enacted but simply to try and “shock” the establishment into adopting more reasonable public policies, then one might even conclude she is trying to further public reason under difficult circumstances.

Our analysis has proceeded using the simple two-person example for analytical clarity. Of course, if these interpretive problems arose between only a handful of persons, they might well be resolved. Suppose Rosie and Jim are close friends who know each other well and meet regularly to discuss politics. Familiarity might in that case overcome the epistemic deficit between them. That deficit would remain, however, between the mass of liberal citizens who do not enjoy face-to-face relations. This returns us to the main contention of

---

chance of one’s vote being decisive and actually affecting legislation, it seems inevitable that citizens will draw upon other motivations when voting.

<sup>394</sup> On such “expressive voting” see Brennan and Lomasky (1997).

<sup>395</sup> Of course, that Rosie wishes to communicate this does not mean that her vote act is interpreted by others in this way. The difficulty of inferring intention from voting behaviour is the very problem we are considering here.

this thesis: that large-scale social coordination cannot presume degrees of mutual knowledge that obtain only in small groups.<sup>396</sup>

#### 4. Outward behaviour Versus Internal Motives

Let us summarise the argument thus far. In the well-ordered society, citizens share the intention of governing their political association on the basis of a reasonable conception of justice. A grouping of persons can relevantly share an intention *iff* each party to the intention knows that all others share the intention. The well-ordered society is thus possible only if citizens can assure one another that each shares the intention to realise a just society. The knowledge that one's fellow citizens have complied with the law does not provide such assurance since such behaviour is coercively enforced. The knowledge that one's fellow citizens have voted on the basis of liberal justice does provide such assurance because such actions are freely undertaken. The question, then, became whether impersonally related liberal citizens could reliably signal to one another their intention to vote on the basis of liberal justice. Two mechanisms for signalling one's intention to their fellow voters were considered, deliberation and the act of voting itself. Both were found wanting. The intentions of our peers in a large-scale society are to a large degree opaque to us. The first desideratum of the well-ordered society cannot be met.

It might be objected that the argument presented here proves too much. If impersonally related liberal citizens are as inscrutable to one another as I have implied, then it seems hard to imagine how any sort of large-scale social cooperation can get off the ground. This includes cooperation in Hayek's Great Society, the members of which need to assure one another of their intention to respect the rules of just conduct. If the arguments presented thus far are sound, then by the same token neither could this Hayekian assurance problem be solved. The remainder of this chapter responds to this objection, which is indeed correct in assuming that members of the Great Society face an assurance problem. However, the epistemic demands it places on them are much less than those placed on liberal citizens tasked with establishing whether or not their peers affirm public reason. Therefore, the prospects for a successful resolution of the "Hayekian" assurance problem seem much better.

---

<sup>396</sup> Rousseau, I believe, was much more sensitive to these epistemic concerns than Rawls. It was one of the reasons the former thought republican government as he understood it would be possible only in small states "where the people can assemble easily and where it's not hard for each citizen to know all the rest" (Rousseau 1968 [1762], 34).

The difference stems from the object of assurance. In the Hayekian great society this is *outward behaviour*. Does this person refrain from using my possessions? Did they keep their promise and pay me on time? And so forth. These behaviours are comparatively easy to perform and recognise, which reduces the interpretive difficulty. Indeed, economising on such interpretive labour is one of the functions of Hayekian rules of conduct, as we saw in chapter one. A foreigner would find it exceedingly hard to join the gift economy of the Kula Ring where interactions are mediated by elaborate and culturally specific understandings. Before they can know what counts as a credible compliance with the elaborate rules, they would have to spend time learning those rules. By contrast, standardised and abstract rules of exchange can often be easily learnt by individuals and communities from a wide variety of backgrounds. It is comparatively easy to learn what the rule consists in and whether one's counterparty has complied with it. Either the goods arrived at the allotted time or they did not. Either the workers received their pay rise or they did not. Thus, the transaction costs to finding out if my counter-parties are committed to the rule or convention in question are kept relatively low.

Crucially, I do not need to know *why* my counter-parties, say, reliably fulfil their part of a contract. Perhaps they are deeply committed libertarians; perhaps they place intrinsic value on obeying the law and contractual law is enforced by the state; perhaps they wish to avoid punishment and there is a credible sanction against breaking contracts in place; perhaps they comply out of sheer habit and tradition. Their souls can remain inscrutable to me. As we have seen, the position of those faced with the legislative assurance problem is exactly the reverse. The object of assurance is not easily identifiable outward behaviour but the *internal motives* underlying such actions. In the various examples we looked at above, Jim could not straightforwardly infer from the fact that Rosie voted for a party or a policy that she had complied with public reason. In each case Jim also needed to know Rosie's internal motivations for voting one way or another.

The Rawlsian might be tempted to say that the distinction between assurance as to outward behaviour and assurance as to internal reasons is ultimately a false one for the following kind of reason. I can only gain reliable assurance that the cooperative outward behaviour of another agent will continue if I have knowledge that they behave in that way for the appropriate reasons. There is a world of difference between knowing that Rosie complies with a rule because it deeply accords with her normative and evaluative standards and knowing that Rosie complies with a rule only contingently when the cost-benefit analysis is

in her favour.<sup>397</sup> The second example suggests that Rosie's outward compliance with the rule is not robust and liable to change as soon as the cost-benefit analysis points towards non-compliance. Only knowledge that Rosie is internally committed to the rule can assure us that her outward compliance is robust and reliable. We shall offer an elaborate critique of this Rawlsian claim in section five of chapter seven, which looks more closely at the question of institutional stability.

For now, we can just say the following. Seeking assurance deeper than outward compliance can just as easily undermine social cooperation as strengthen it. Consider how actually-existing liberal democracies contain within them many individuals whose deepest normative ideals may not be liberal and indeed radically diverge from liberal norms. Actually-existing liberal democracies contain communists, ethnic nationalists, communitarians and deep greens among others. The great majority of these persons, the great majority of the time, comply with the law. They respect persons and property, pay their taxes, comply with state officials and so on. They do so despite the existing institutional structure being inconsistent with their first-best ideals. One reason why they might do so is that *there is no expectation on their part of such a consistency*. In other words, they take an implicitly Augustinian approach to political and social institutions. They take as normal a disjuncture between their deepest normative ideals and the political institutions under which they live.<sup>398</sup> Enter the political philosopher who tells them it is right and proper to expect the political and legal institutions that coerce them to accord with their most deeply held normative aspirations. The philosopher's eloquence persuades them to discard the Augustinian realism they previously held implicitly. The problem, however, is that a liberal society evidently cannot meet these newly ignited normative demands. The basic tenets of, say, communism or ethnic nationalism are incompatible both with one another and the basic tenets of a liberal order.

---

<sup>397</sup> On this see Rawls' distinction between a *modus vivendi* and what he sees as the more normatively secure "overlapping consensus" in Rawls (1996, chap. 4).

<sup>398</sup> I here follow Jacob Levy's (2017) secular reinterpretation of Augustine's political theology. On Levy's interpretation, an Augustinian approach to politics means that one accepts that the ideal political system – the "City of God" – is unavailable to us in the profane and sinful "City of Man". In the latter one has to make do with a political settlement that falls far short of the City of God but still has enough worth to merit one's allegiance. Only in the afterlife can the ideal of the City of God be attained. Levy argues that one can adopt a secular version of this Augustinian political attitude if they substitute Augustine's City of God for their own normative ideal and set aside the possibility of arriving at their ideal in the afterlife. I think many liberal citizens with distinctly non-liberal political views adopt something like this approach to their political institutions. Though I write that they do so "implicitly" because they are unlikely to express this sentiment in the explicit terms of Augustine's political philosophy.



The *generation* of such *unrealisable* expectations could itself be a recipe for institutional stability.

## 5. The Necessity of Population-Level Judgements

Another objection to the argument might run as follows.<sup>399</sup> The objection would question a key premise of the argument presented thus far, namely, that a *population-level judgement*, as to what proportion of the nation as a whole is committed to liberal justice, is required. On the basis of this assumption we have inferred that Rawlsian citizens cannot assure one another of their commitment to liberal justice whether via the mechanism of deliberation or voting. The objection would argue that our argument has exaggerated the epistemic demands involved. It is not true that each citizen needs to make such a population-level judgement. They can simply take stock of the people around them, of whom they do have direct or intimate knowledge, and then render their decision to comply with public reason contingent on assurance from this much more tractable local sample. By analogy, this is how social norms operate. As Christina Bicchieri has shown, mundane and everyday social norms such as promise-keeping have the same characteristics of Rawls' assurance problem. On her account, a social norm holds in a network if each member (a) believes all other members follow the norm; (b) believes all other members expect them personally to follow the norm and (c) compliance with the norm at least sometimes goes against self-interest narrowly understood.<sup>400</sup> Compliance with public reason can be understood as a social norm in Bicchieri's sense. It clashes with self-interest per (c). Compliance with it is a legitimate expectation we can have of others per (b). The compliance of each member is contingent on their belief that others reciprocate per (a).

Consider how the norm of promise-keeping takes hold in a population as large as the United Kingdom. No one does or indeed could gauge what proportion of the total UK population complies with the norm. Rather, each just keeps track of whether those they interact with keep their promises. If enough do, they keep their promises. Crucially, one tends to have enough concrete knowledge of their immediate peers to know whether they comply or not. The problem of opacity does not arise as it does with population-level judgements. As long as each party in even a very large network gains this more tractable, local form of assurance, the norm of promise-keeping can become entrenched among a high number of people

---

<sup>399</sup> I thank Temi Ogunye for advancing this objection.

<sup>400</sup> (Bicchieri 2006, chap. 1)

simply due to network affects. The objector would simply ask why Rawlsian assurance could not operate on the same mechanism. Each individual citizen just needs to know that the voters in her neighbourhood or work place or family, of whom she has concrete knowledge, exercise their right to vote in accordance with public reason. If enough citizens gain this local form of assurance the commitment to public reason too could spread through such network affects. The objection provides a plausible account of how most norms and conventions, including Hayekian rules of just conduct, might spread through a large population. It is mistaken, however, in assuming that this local form of assurance is adequate for the entrenchment of the norm of public reason. For when an agent is seeking assurance as to how the mass of their fellow citizens will use their right to vote, they necessarily have to make population-level judgements.

A thought experiment might illustrate why. Let us further consider the norm of promise-keeping. If I may be allowed yet more uncreative licence, imagine a fictional country with two small towns – Trustinium and Untrustinium. As one might guess, promise-keeping is a strong norm in the former but not the latter. Suppose I live in Trustinium and I have learnt to keep my promises because those around me generally do so. This seems rational because for the most part I only deal with fellow residents of Trustinium. When it comes to promise-keeping, all I need to know is whether those *I actually interact with or am likely to interact with follow the norm*. I do not need to make population-level judgements since I do not regularly interact with each and every single other member of the population. Since I rarely have interactions with members of Untrustinium, for instance, it is of little consequence to me that they are untrustworthy. I can safely limit the scope of my concern to my local circumstances. The equivalent can be said for a resident of Untrustinium who has learnt not to keep her promises. This seems rational given who she interacts with. The fact that people might be trustworthy elsewhere is not relevant.

Now let us now consider the same scenario save for the norm requiring assurance being that of voting according to public reason. Citizens of Trustinium are committed Rawlsians while Untrustinium is full of defectors from Rawlsian justice. The two scenarios are disanalogous in this sense: when the assurance model is about the legislative power of the state that is held equally with all citizens, population-level judgements *are* required. The scope of my concern would now rationally extend to how *all* of my fellow voters behaved. They each have an equal sovereign stake in political power and so unlike the promise-keeping case, the behaviour of all directly affects me. I cannot rationally limit my concern to whether those I regularly interact with comply with public reason. If I am a resident of Trustinium and I take

my local experience as representative of the whole population, I will likely overstate the latter's commitment to public reason. Conversely, if I am a resident of Untrustinium and I do the equivalent, I will likely understate their commitment to public reason. By necessity I have to cast a wider net to gain the assurance I seek. And so for this reason, population-level judgements, while not necessary for the assurance apropos of most norms, including Hayekian rules of just conduct, are necessary in the Rawlsian model.

## **Conclusion**

Even if we grant Rawls' the idealised assumption that all citizens affirm public reason, they cannot credibly signal this fact to one another. Perhaps, however, the difficulty is that Rawls asks too little of liberal citizens. Collective political agency doesn't simply require citizens to show their commitment to liberal justice in the voting booth. It requires them also to engage in regular and sustained political deliberation on the common good. Only then can citizens can form a genuine consensus as to how their political association should be governed. We already cast some doubt on this claim earlier in this chapter. We observed that deliberation performs poorly as a signalling mechanism in solving Rawls' assurance problem. In the next chapter we will develop further our critique of deliberative democracy.

## Chapter Five: Aspirational Associationalism – Deliberative Democracy

### Abstract

This chapter considers whether collective political agency would be possible in a society characterised by widespread subscription to deliberative democratic practices. It makes three arguments. First, the very conditions that deliberative democrats deem necessary for collective will-formation are possible only in face-to-face settings. Second, various attempts to scale up deliberative democracy fail to solve this problem. At most they will lead to a proliferation of “mini-publics” each of whom *internally* meet the necessary conditions for collective will-formation. However, these “collective wills” will be shared only by members of the mini-public in question and not society as a whole. Third, it draws on Alexis De Tocqueville to argue that the tendency to centralise and collectivise decision-making undermines those very civil society associations in which the deliberative ideal might be possible.

### Introduction

Deliberative democrats are all too aware that citizens of actually-existing liberal democracies do not exercise meaningful collective agency. However, they propose a more robust commitment to deliberation as the antidote to this problem. It is hoped that such deliberation will lead to genuine collective will-formation in contrast to the status quo. James Fishkin and Bruce Ackerman propose a national “Deliberation Day” to rectify the apathy and ignorance of the mass of regular voters.<sup>401</sup> Amy Gutmann and Dennis Thompson argue that Rawls should adopt a more robust commitment to deliberative democracy because otherwise the consensus of liberal citizens would be limited to only the abstract content of justice as fairness. It would not extend to the concrete details of its application.<sup>402</sup> Joshua Cohen argues that widespread deliberation will encourage in citizens an attachment to the common good, unlike actually-existing liberal democratic practice in which citizens vote on the basis of their private and self-centred concerns.<sup>403</sup> Jürgen Habermas has long seen his theory of discursive will-formation as completing the promise of popular sovereignty that has been only partially realised since the French Revolution.<sup>404</sup> Rawls himself had an ambivalent attitude towards deliberative democracy. He was famously criticised by Habermas for his “monological” original position which lacked any real deliberation.<sup>405</sup> As we

---

<sup>401</sup> (Ackerman and Fishkin 2004)

<sup>402</sup> (Gutmann and Thompson 1998)

<sup>403</sup> (Cohen 1997)

<sup>404</sup> E.g. See “Popular Sovereignty as Procedure” in Habermas (1996, 463-491).

<sup>405</sup> (Jurgen Habermas 1995)

noted in the previous chapter, he sometimes says it is unrealistic to expect most citizens to dedicate substantive portions of their lives to political deliberation.<sup>406</sup> In other places, however, he explicitly describes the well-ordered society as a form of deliberative democracy and suggests that the aim of public reason is to encourage as much deliberation as might be possible, citing favourably his former student Joshua Cohen.<sup>407</sup>

This chapter critiques deliberative democracy on the grounds that it does not scale up in a large society. The argument to follow is divided into five sections. The first outlines deliberative democracy's scale problem. According to deliberative theorists such as Habermas and Cohen, the epistemic presuppositions for genuine collective will-formation are met only in small face-to-face groupings. Sections 2-4 consider and critique three attempts to overcome the scale problem. Some think we can treat the deliberative conclusions arrived at by a group of persons randomly chosen from the population-at-large as representative of the conclusions the latter grouping would arrive at were they all to somehow deliberate together. Section two argues that this representation thesis lacks foundation. Ackerman and Fishkin propose "Deliberation Day," an official public holiday taking place before a major national election and during which eligible voters are given the opportunity to engage in political deliberation. Section three argues that, at most, this might lead to better informed voters but not to the formation of a society-wide collective will. Section four considers "systemic" approaches to deliberative democracy. These approaches acknowledge that a large-scale society will inevitably consist of pluralistic and polycentric deliberative fora and try to re-formulate deliberative theory to be more realistic as a result. It is argued that such approaches are in the end difficult to distinguish from the aggregative status quo that deliberative democracy is supposed to replace. Section five takes a slightly different tack. It argues that deliberative democrats lack a robust account of the conditions under which agents will have sufficient knowledge and motivation to undertake deliberation. Drawing on Alexis De Tocqueville in particular, it argues that these conditions are best satisfied in a decentralised social order where deliberation can aid communities in solving

---

<sup>406</sup> E.g. (Rawls 2001, 143)

<sup>407</sup> (Rawls 1997, 772–73); still, Rawls (1997, 769) also thinks the responsibility to use public reason in deliberation falls primarily on politicians and judicial officials rather than regular citizens, on whom falls the responsibility of voting according to public reason. His overtures to deliberative democracy notwithstanding, this suggests that he envisages much less deliberation by regular citizens in his ideal regime than Habermas and Cohen do in theirs.

local problems. Extensive welfare guarantees, of the sort often favoured by deliberative democrats, may undermine the conditions in which the deliberative virtues flourish.

### **1. Small-Scale Limitations of Deliberative Democracy**

Deliberative democracy aims to rectify a defect of aggregative democratic procedures. These treat voter preferences as exogenously given inputs and use them to generate a collective “decision” using an aggregation rule, most typically a majoritarian function. For deliberative democrats, the problem with aggregative democracy is not simply that voters’ pre-existing preferences might be ill-formed, prejudiced or even non-existent.<sup>408</sup> It is that the “output” of such aggregation rules often fall short of a coherent or consistent general will, as evinced by the familiar aggregation problems that we covered in chapter three. The confused “outputs” of these democratic procedures reflect the inputs received from voters. Under aggregative democracy, voters act on their disparate and idiosyncratic preferences rather than a shared understanding of the common good. For the deliberative democrat, it is little surprise then that existing democratic practice does not always reveal the people to have a rational will. Deliberative democracy, on the other hand, emphasizes the importance of building deliberative consensus before voting. Democratic deliberation should follow procedures that respect the freedom and equality of all participants. It would not take voter preferences as given exogenously to the democratic process but offer voters the opportunity to change and revise them. The deliberative process of revising voters’ initial preferences should also facilitate a convergence of their views towards a rational consensus. This consensus would then inform their voting decisions.<sup>409</sup>

Deliberative theorists have identified necessary conditions for the formation of such a consensus. Joshua Cohen’s account of the formal presuppositions of deliberative democracy is representative of the literature.<sup>410</sup> These suppositions are four in number:

1. The participants see themselves as bound only by the results of their deliberation and treat as action-guiding the results of the deliberation.
2. The participants back-up their views with reasons.
3. Each participant has an equal voice in the deliberation.
4. The participants aim for a “rationally-motivated consensus”.

---

<sup>408</sup> The last was Phillip Converse’s (2006) famous claim.

<sup>409</sup> (Cohen 1997); (Jürgen Habermas 1996, 2001); (Gutmann and Thompson 1998); (Mansbridge et. al. 2012)

<sup>410</sup> (Cohen 1997, 74–75) c.f. Habermas (1983)

The first condition invokes the kind of stipulations on discourse that one might find in a jury setting. The second condition stipulates that deliberation should be reasoned. It insists that only reason, not force or manipulation, can be used to gain support for a proposal. Cohen explicitly cites Habermas' requirement that "no force except that of the better argument is exercised".<sup>411</sup> The reasons should be formulated in such a way as to be acceptable to the other participants given the "diverse preferences, convictions, and ideals concerning the conduct of their own lives".<sup>412</sup> The third condition of an equal voice in deliberation means that "[each] can put issues on the agenda, propose solutions, and offer reasons in support of or in criticism of proposals. And each has an equal voice in the decision".<sup>413</sup> The fourth condition of a rationally-motivated consensus is achieved when reasons are found that are persuasive to all deliberators in the group.

These very demanding criteria could be reasonably approximated only in small groups characterised by face-to-face contact. Supposing the deliberative process is to be at all a practical affair, then the third requirement of equal voice alone seems to dictate small numbers. When considering Cohen's proposal, Michael Huemer notes that the US population is over 300 million. He asks sardonically: "[what] would society be like if every one of those individuals could call a press conference to discuss their latest policy idea?"<sup>414</sup> The epistemic hurdles to meeting these demanding discursive conditions on a large-scale do not end there. If I am to propose reasons that are acceptable to all the other deliberators, given their diverse "preferences, convictions, and ideals," I need to be cognisant of what these might be. I might be able to get some traction on the personal ideals of my fellow deliberators if we were a handful of persons. I will not be able to get any traction on such ideals if my fellow deliberators are, say, 327 million fellow citizens of the United States. If I am to enjoy a *rationally-motivated* consensus with my deliberative peers, it is insufficient to merely convince them of my proposal. I also need to know that they are in *reasoned agreement* with it. That is, I need to know that all others are in agreement for the right reasons that I have offered them. I have not arrived at a rationally-motivated consensus with them if they have simply misunderstood the nature of my proposal or agreed for inappropriate reasons.<sup>415</sup> Nor is knowing that they have accepted my proposal for the right

---

<sup>411</sup> (Cohen 1997, 75)

<sup>412</sup> (Cohen 1997, 72)

<sup>413</sup> (Cohen 1997, 74)

<sup>414</sup> (Huemer 2013, 63)

<sup>415</sup> Perhaps, for instance, they are expeditiously agreeing with me so they can go home after a long, tiring day of deliberation. Having performed jury service in the UK, I offer this as more than a hypothetical scenario.

reasons sufficient. They have to know that I know this. In turn, I need to know that they know *that—ad infinitum*. It is implausible to think that such thick intersubjective understanding can be gained amongst any but face-to-face deliberative groupings.

Cohen's requirements for deliberative will-formation, then, are similar to Bratman's requirements for the formation of a joint-intentional action. Bratman outlines the conditions under which a group shares an intention to act. A group shares an intention that they perform an act if each member intends that they perform that act, intends to do their part in the performance and knows the equivalent to be true of all other members of the group. Cohen outlines the conditions under which a group shares an intention to act for the *same reasons*. A group shares a *rationaly-motivated* intention that they perform an act if each intends the act for a reason or set of reasons accepted by all in the group and knows the equivalent to be true of all other members of the group.<sup>416</sup> The demanding mutual knowledge requirements of both limit their applicability to small, face-to-face groups. It follows that the conditions necessary for genuine deliberative will-formation are readily met only in miniscule subsets of the population of a large-scale society. Thus, even if all members of such a society were highly motivated to engage in regular political deliberation, it would be insufficient to form a society-wide deliberative consensus. The most we can expect, then, from such widespread adoption of deliberative practices is the creation of a mass of mini-publics each of which *internally* meets the requirements for deliberative will-formation. Deliberative theorists are more than aware of this scale problem and have made various proposals to overcome it. The sections to follow will review three such proposals and will argue all fail to solve the scale problem. At most, they might help voters to be slightly more informed about politics, but none will lead to a large-scale deliberative consensus.

## 2. Scaling Up Deliberative Democracy – The Representation Approach

The first proposal to be considered is Robert Dahl's famous proposal for a "minipopulus".<sup>417</sup> Dahl proposed that a thousand citizens be selected randomly from the entire population. This grouping would deliberate on an issue for an extended period of time and deliver its verdict. Dahl claims that their verdict would

---

<sup>416</sup> The difference is that on Bratman's account, a group could share an intention to perform an act but each individual member of the group might have different reasons for performing it.

<sup>417</sup> (Dahl 1991, 238–41)



... “represent” the judgment of the demos [as a whole]. Its verdict would be the verdict of the demos itself, if the demos were able to take advantage of the best available knowledge to decide what policies were most likely to achieve the ends it sought. The judgments of the minipopulus would thus derive their authority from the legitimacy of democracy.<sup>418</sup>

In more recent years, James Fishkin has advanced a close variant of Dahl’s idea. Fishkin has tested deliberative democracy empirically by organising deliberative groupings for experimental purposes.<sup>419</sup> Each experiment consists of random and statistically representative samples of 100-200 citizens tasked with systematically deliberating a public policy matter. The pre- and post-deliberation opinions of the participants are polled. Fishkin has conducted such experiments all around the world. He finds that citizen opinions often change to a large degree throughout the process and that they are better informed at the end of the process compared with the beginning.<sup>420</sup> His experimental work constitutes some evidence that deliberative democracy at the small, face-to-face level works as deliberative theorists predict. Dahl also thinks, like Fishkin, that the conclusions arrived at by these small-scale groupings somehow track wider public opinion. He calls these experiments “deliberative polls”. In contrast to normal polling, which ascertains the untutored opinions of a sample of the public, deliberative polling ascertains the considered opinion of a sample that have deliberated the matter at hand with one another. The hope is that their conclusions “represent what the public [as a whole] would think about the issue if it were motivated to become more informed and to consider competing arguments”.<sup>421</sup>

In evaluating Dahl and Fishkin’s proposal, we need to distinguish deliberative polling from “lottocracy”. The latter is a term coined by Alexander Guerrero for a system whereby political officials are chosen by lot rather than elected. Guerrero argues that lottocracy is preferable to elective systems on the ground of several considerations. These include it being less susceptible to corruption and selecting for officials with longer time horizons than elected ones.<sup>422</sup> If lottocracy is a form of governance, deliberative polling is the utilisation of lottocracy for one particular objective, namely, to arrive at a decision representative of public opinion at large. It is this particular rationale for lottocracy that is subjected to critique in this section.<sup>423</sup> Neither Fishkin nor Dahl, to my knowledge, provides any reason why the

---

<sup>418</sup> (Dahl 1991, 340)

<sup>419</sup> (Fishkin 2003)

<sup>420</sup> (Fishkin 2003, 129–30)

<sup>421</sup> (Fishkin 2003, 129)

<sup>422</sup> (Guerrero 2014, 165–70)

<sup>423</sup> Thus leaving open the possibility that lottocracy might still be desirable on some of the other grounds listed by Guerrero. This chapter is agnostic on that wider question.

conclusions of a group selected by lot would represent the considered opinion of the whole population. One possibility is that they subscribe to a convergence thesis about human reasoning. It would run as follows:

If a deliberative group meets or comes reasonably close to meeting appropriate deliberative conditions – e.g. conditions (1)-(4) above – then the conclusion they reach with respect to a particular proposal – i.e. whether to accept or reject it – and the reason on which they reach that conclusion will be the same as those reached by any other deliberative group meeting or coming reasonably close to meeting those same conditions.

In other words, the convergence thesis states that the free use of reason by different deliberative groupings would lead to the same conclusion. If so, the conclusion reached by members of a randomly selected group would track that reached by the population as a whole if the latter could all somehow deliberate with one another.

Now, one would not want to dismiss the convergence thesis *tout court*. There might well be some conclusions that we would expect (or at least hope) that any deliberative group would reach if they considered the matter-at-hand under appropriate conditions. One would hope that demonstrably unwarranted beliefs – flat-earth theory for instance – will be rejected by any competent group of deliberators. Still, if we accept even a mild dose of Rawls' reasonable pluralism, the convergence thesis will not be true of a whole host of proposals.<sup>424</sup> These are proposals the evaluation of which is subject to burdens of judgement. The evidence might be complex and difficult to evaluate, their evaluation might be contingent on multiple considerations the weighing of which might not be obvious, there might be powerful normative considerations both in favour and against the proposal, and so on.<sup>425</sup> Following Rawls' cue, we might say that each person will judge such proposals according to his or her personal *evaluative standards*.<sup>426</sup> These differing evaluative standards mean that competent reasoners will non-culpably come to divergent conclusions when deliberating complex subject matter.<sup>427</sup>

This Rawlsian thesis was originally intended to apply to *individual* deliberation. We can, however, straightforwardly apply it to deliberative groups. The conclusions arrived at by the

---

<sup>424</sup> Recall that section three of chapter one did not deny that the free use of human reason leads to divergent conclusions. The claim was that we have limited knowledge as to what our peers think in an impersonal society. This ignorance greatly limits the extent to which we can be described as being *in a state* of agreement or disagreement with them.

<sup>425</sup> (Rawls 1996, 56–57)

<sup>426</sup> Though I take the term from Gaus (2012a, 276).

<sup>427</sup> I write that they “come to divergent conclusions” rather than “disagree” for the reasons outlined in section three of chapter one.

members of a deliberative group will be contingent on its members' particular array of evaluative standards. Imagine two deliberative groups considering abortion law. Group A contains only deliberators with conservative Christian beliefs. Group B only atheists. If both groups were to settle on a conclusion and a reason for that conclusion, it seems highly unlikely that they would be the same. Common sense suggests A will arrive at a much more restrictive policy than B. Deliberative democrats often hope that the educative function of deliberation will narrow down differences in evaluative standards. Even if this were the case, it would be of little use in rescuing the convergence thesis here. While Deliberation might narrow evaluative standards *within* group A or B, there is no guarantee it will narrow the differences *between* them. Indeed, it seems just as possible, if not more so, that it would exacerbate the latter differences. *Post*-deliberation, members of A could have converged in their evaluative standards such that the average member of A now adopts a *more* restrictive stance towards abortion than they did pre-deliberation. Likewise but conversely, the average member of B would adopt an even *less* restrictive stance. Where the education function takes the members of the deliberative group depends on where those members begin. There is no guarantee that the final destination is the same for all groups given their cognitive diversity. The convergence thesis, then, fails to substantiate the representation approach.

Perhaps the representation claim is grounded on another thesis. Fishkin's deliberative polls are based on a statistically representative sample. If so, Fishkin might affirm something like the following:

If a statistically representative sample of a larger grouping arrives at a rationally-motivated consensus with respect to a particular proposal(s), that consensus is either the same as or otherwise reasonably tracks the rationally-motivated consensus the larger grouping would arrive at if its members were all capable of deliberating together.

We can read this view as a modification of the convergence claim. It does not claim that the conclusions of a group selected purely by lot will track those of the demos as a whole. It claims that the conclusions of a statistically representative sample would do so. This modified form of the convergence claim, however, still seems mistaken, particularly for a large and diverse population. It falls foul of the fallacy of composition. The sampled categories will internally admit of great variation. For example, 3.3% of the British population is black. It hardly follows that the opinions of three black Britons chosen at random for a deliberative poll of, say, 100 people will adequately track the diversity of opinion in the British black community. Even if they somehow did, the whole point of deliberation is that it might lead

to the revision of one's initial views. If after deliberation these three individuals change their opinion in a certain manner, it would surely be presumptuous to assume that other members of the black community would have also shifted their views in this manner had they deliberated the proposal under discussion. The modified convergence thesis also fails to substantiate the representation approach. We conclude, then, that the representation approach fails to scale up deliberative democracy.

### 3. Scaling Up Deliberative Democracy – Deliberation Day

Along with Bruce Ackerman, Fishkin has another proposal called “deliberation day”. I here draw on their book of the same name.<sup>428</sup> Unlike deliberative polling, this proposal involves participation on a mass scale. The authors propose a very concrete and detailed institutional proposal for how deliberation day may be implemented in the United States. It will be a national holiday held two weeks before every presidential election. Despite the name it will actually be a two-day affair. All citizens are eligible to take one day off to participate for which they are paid an honorarium.<sup>429</sup> Participation is voluntary. Deliberation day events will happen all across the country. On the day, there will be two kinds of deliberative formats. In the morning, the participants first convene in small groups of 15 and then convene for a second set of larger debates in “citizens’ assemblies” of 500 persons. The participants in the first debates formulate and vote on a list of questions to be put forth to representatives of the presidential candidates in the second debates, during which they listen to the representatives’ answers. This process is repeated in the afternoon.<sup>430</sup>

Ackerman and Fishkin are distressed by the apathy and ignorance of the average voter, which renders the democratic process vulnerable to capture by special interests. Their proposal is intended to show that talk of popular sovereignty is not “hot air”. Citizens “have the fundamental right to determine the broad direction of public policy through electoral decisions made on the basis of popular deliberation”.<sup>431</sup> Deliberation day is offered as the institutional mechanism through which such popular sovereignty might be exercised. The hope is that it will generate a substantive and informed public opinion that can hold presidential candidates to account. As before, in our assessment we should distinguish

---

<sup>428</sup> *Deliberation Day* (Ackerman and Fishkin 2004)

<sup>429</sup> (Ackerman and Fishkin 2004, 3)

<sup>430</sup> (Ackerman and Fishkin 2004, 27–37)

<sup>431</sup> (Ackerman and Fishkin 2004, 11–12)

between the all-things-considered case for an institution and a particular rationale. Perhaps deliberation day is a good idea because it would, say, improve voter knowledge of basic political facts.<sup>432</sup> This chapter takes no position either way on that empirical question. Our question is whether deliberation day will facilitate the emergence of a large-scale deliberative consensus that can, in the authors' own words, "determine the broad direction of public policy". The answer is negative. It is hard to see how any such consensus would arise from these debates. At most, the small groupings of 15 persons might internally converge on a considered group opinion with respect to some issues.<sup>433</sup> These agreements, however, would not extend beyond their small groupings. So we still have just a plethora of mini-publics at best. Deliberation day also fails to scale up deliberative democracy.

#### **4. Scaling Up Deliberative Democracy – Systemic Approaches**

In recent years, deliberative democrats have shifted towards so-called systemic approaches.<sup>434</sup> Systemic deliberative democrats aspire towards sociological realism. They recognise that actually-existing liberal democracies have multiple and distributed centres of democratic deliberation and contestation, ranging from the courts to various legislative chambers to civil society organisations. In the terminology of the literature, democratic authority in large and complex societies is "decentered".<sup>435</sup> Rather than conceive of a singular public whose will is expressed by a legislative centre, we should model a pluralistic and differentiated public sphere consisting of many "nodes" of deliberation.<sup>436</sup> The focus should not be on face-to-face deliberation in just one arena but on providing an account of how these various deliberative fora interact with one another. The systemic approach is often presented as an explicit solution to deliberative democracy's scale problem. This section argues it is not. Rather than showing how deliberative will-formation is possible on a large-scale, systemic approaches are an implicit concession that the scale problem can't be solved. Such theories look little different from actually-existing democratic practice, the radical reform of which was the original aim of deliberative theory. For once we recognise that

---

<sup>432</sup> Though it would do little to rectify the more radical forms of voter ignorance we identified in chapter three.

<sup>433</sup> It is only in these smaller groupings that the participants actually deliberate with one another, in order to formulate questions for the larger citizen assemblies. In the latter, they only listen to representatives of the presidential candidates.

<sup>434</sup> (Habermas 1996); (Mansbridge et. al. 2012); (Bohman 2005); (Dryzek 2010)

<sup>435</sup> (Bohman 2005, 296–97); (Habermas 1996, 298–302)

<sup>436</sup> (Mansbridge et. al. 2012, 10)

democratic opinion is pluralistic and decentred, there is no reason to expect the multiple voices of the public sphere to cohere into any deliberative consensus. We will show how this is the case with respect to Habermas' version of the systemic approach, as outlined in his *Between Facts and Norms*.<sup>437</sup>

Habermas distinguishes between an "anarchic", "weak" and "wild" public sphere that operates in the periphery of society and the formal legislative and judicial centre – a "subsystem specialized for collectively binding decisions".<sup>438</sup> He proposes a model of deliberative politics based on a division of labour between these two spheres. Only the political centre is capable of action in the sense of promulgating coercively-binding laws. However, reasons for action are supplied by public opinion generated in a deliberative but polycentric public sphere. The model is explicitly presented as a third way between a more communitarian model of deliberative democracy in which the state is seen as a vehicle for the undifferentiated and unified common will of the populace, and a liberal model emphasising the institutionally mediated form that the expression of people's sovereignty need take in a complex society.<sup>439</sup> In his own words, his theory "depends not on a collectively acting citizenry but on the institutionalisation of ... procedures and conditions of communication".<sup>440</sup>

It is difficult to properly evaluate Habermas' proposal. His arguments are pitched at a high level of abstraction and it is not quite clear what the interaction between the anarchic public sphere and the centre is supposed to look like. He does offer suggestive metaphors. The public sphere "programs" or "supplies" the centre or otherwise provides it with "political potentials".<sup>441</sup> It acts as a "sounding board for problems" or "warning system".<sup>442</sup> It seems that communicating new ideas or concerns to the legislative centre is one of the public sphere's core functions. Habermas refers positively, for instance, to the effectiveness of civil society associations in bringing a range of issues into public consciousness, ranging from the problems of nuclear proliferation to the demands of the feminist movement.<sup>443</sup> He also emphasises repeatedly the dynamic, open-ended and polycentric nature of the public sphere. It consists of "an open and inclusive network of overlapping, subcultural publics

---

<sup>437</sup> (Habermas 1996)

<sup>438</sup> (Habermas 1996, 307, 300)

<sup>439</sup> (Habermas 1996, 297–98)

<sup>440</sup> (Habermas 1996, 298)

<sup>441</sup> (Habermas 1996, 300, 352, 363)

<sup>442</sup> (Habermas 1996, 359)

<sup>443</sup> (Habermas 1996, 381)

having fluid temporal, social, and substantive boundaries”.<sup>444</sup> It also “represents a highly complex network that branches out into a multitude of overlapping international, national, regional, local, and subcultural arenas”.<sup>445</sup>

Habermas’ notion of the public sphere, then, appears to have a strong family resemblance to Hayek’s idea of an emergent social order.<sup>446</sup> Indeed, Habermas sounds positively Hayekian when he says that the public cannot occupy “*the position* of a macrosubject supposed to bring society as a whole under control and simultaneously act for it”.<sup>447</sup> The public opinion emanating from this public sphere is not a singular, unitary or consistent phenomenon. Instead, we can see the public sphere as a network through which information is transmitted between very many focal points - “the streams of communication are, in the process, filtered and synthesized in such a way that they coalesce into *bundles of topically specified public opinions*”.<sup>448</sup> The question, then, is how these diverse public opinions interact with the formal law-making centre. Drawing on Bernard Peters’ work, Habermas offers the metaphor of “sluices” through which these streams of public opinion are transmitted to the centre.<sup>449</sup> However, if we take this metaphor seriously, the resulting picture is hard to distinguish from the relationship between legislative law-making bodies and wider civil society that we already find in large-scale democracies. It is quite clear that the political centre is going to receive a cacophony of discrete and “topically specified public opinions,” many of which will issue incompatible directives. Certain familiar difficulties crop up. The ideas emanating from the public sphere that meet Habermas’ approval – ranging from nuclear disarmament to environmentalism to feminism – tend toward the “left-” or “progressive-” side of the political spectrum. We can imagine rather different ideas stemming from more “conservative” parties of the public sphere. Which among these and many other voices, which will often be in tension with one another, should the legislative centre heed? It is hard to avoid the conclusion that those with access to political power will listen to some amongst the cacophony of voices but also ignore very many others.

---

<sup>444</sup> (Habermas 1996, 307)

<sup>445</sup> (Habermas 1996, 373)

<sup>446</sup> Hayek (1982c, 140–41) also offers a description of society as a complex network of relations in the

third volume of his *Law, Legislation and Liberty*.

<sup>447</sup> (Habermas 1996, 372); original emphasis.

<sup>448</sup> (Habermas 1996, 360); emphasis added.

<sup>449</sup> (Habermas 1996, 300)

We need not even posit self-interested motives amongst the political class to reach this conclusion.<sup>450</sup> Given the quantity, diversity and frequently contradictory nature of the “topically specified opinions” in the public sphere, even the most virtuous and impartial of political officials would be unable to learn about and act upon all of them. They would be forced to exercise discretion in selecting those among the many “public opinions” with which they are familiar and/or find important. As a result, many parts of the public sphere will find their views and opinions overlooked by the political class. This is a phenomena already prevalent in established democratic practice. Carolyn Hendriks has documented an example of it in some empirical research. Her case study is of a citizens’ conference held in Germany on the topic of “genetic diagnostics”.<sup>451</sup> The conference had a mixture of formal discussions dominated by experts and more informal ones open to participation from the wider public. She observes that against the “hyperactive discursive setting” of the large-scale public sphere, “the citizens’ conference struggled to gain any attention in the political and public domain”.<sup>452</sup> The fact that many if not most citizens do not exercise a meaningful say in the formulation of democratic laws is what motivated the deliberative turn in the first instance. This fact appears unchanged in Habermas’ model of the relationship between a deliberative yet highly fragmented public sphere and a unitary legislative centre forced to select which bits of the public sphere it should listen to. Habermas’ theory purchases sociological realism but only by demonstrating the infeasibility of a society-wide deliberative consensus.

Some other advocates of the systemic approach have criticised Habermas for his dichotomy between centralised legislative authority at the core and the anarchic and pluralistic public sphere at the periphery. They argue that *decision-making* authority is just as decentred and polycentric as the sources of *public opinion*.<sup>453</sup> They might object, then, that their construal of deliberative systems is not vulnerable to the critique just presented, since it does not posit one monocentric pole of decision-making that all the heterogeneous voices of the polycentric public sphere try to influence. Thus Mansbridge et al argue that our understanding of deliberative systems should extend to include both informal centres of decision-making in civil society and formal centres of decision-making in international and transnational institutions.<sup>454</sup> They then claim that a rational deliberative consensus, at least

---

<sup>450</sup> Such that politicians would heed only those amongst the many deliberative voices that would help them win elections or who emanate from agents who fund their campaigns.

<sup>451</sup> That is, how public policy should respond to the prospective emergence of gene-editing technology.

<sup>452</sup> (Hendriks 2006, 498)

<sup>453</sup> This is similar to the diverse model of rule-making authority that I defended in chapter three.

<sup>454</sup> (Mansbridge et. al. 2012, 9)



with respect to some issues, can be seen as an outcome emerging from the interaction of multiple discursive and decision-making nodes. They call these outcomes “societal choices” or “societal conclusions”. For instance, they write:

... The widespread societal conclusion that discrimination in hiring by race and gender is unjust is reasonably described as a collective decision, resulting in part from certain binding state decisions but also in large part from hundreds of millions of individual and institutional decisions based on widespread collective discussion and interaction.<sup>455</sup>

I believe there to be an error in this passage similar to that we identified in chapter two with Gaus’ *The Order of Public Reason*. It ascribes too much intentionality to emergent outcomes. The passage implies that all these individuals and institutions consciously deliberated on the rights and wrongs of discrimination, concluded it was wrong, and then changed their practices. Steadily the argument against discrimination spread across the social space until all become convinced of its validity. Now, no doubt some such direct persuasion and argument did play a role in the entrenchment of these anti-discrimination norms. But another significant factor leading different individuals and institutions to adopt anti-discrimination practices are the dynamics of social norm-following. Per Christina Bicchieri’s aforementioned account, norms are default rules of behaviour to which we conform because we believe that a sufficiently large number of our peers follow the norm and that a sufficiently large number of them *expect* us to follow it.<sup>456</sup> Bicchieri’s account implies a complex relationship between the norms that members of a group follow and their personal beliefs. A norm of which no member of a group *personally* approves could still be unanimously subscribed to if each believes sufficient other members of the group approve the norm and expect them to follow it.

This is an extreme case, but it highlights how norms can spread spontaneously without anyone strongly approving or disapproving of them but simply through the power of mutual expectations in a large network. Once the emergent norm against discrimination hits critical mass, even those who have doubts about it have incentives to comply.<sup>457</sup> Of course, personal normative beliefs are not *irrelevant* to the emergence of norms. One can reject a norm or try entrepreneurially to reform it if they feel strongly enough about the matter. But the spread of a norm is never solely dependent on individuals being *personally* satisfied as to its value. Each agent (individual or institutional) has to consider both their personal moral attitude to

---

<sup>455</sup> (Mansbridge et. al. 2012, 8)

<sup>456</sup> (Bicchieri 2006, 5)

<sup>457</sup> Much, as we saw, was the same with the emergence of an equilibrium rule in Gaus’ Kantian coordination game.

the norm and also what they believe to be the expectations of other agents in their networks. Thus, I might consider a norm much less than ideal but follow it because I believe I am expected to.<sup>458</sup> If this analysis is sound, we cannot straightforwardly infer widespread ethical agreement on the worthiness of a norm from the empirical fact of widespread conformity to it. As with Gaus' Kantian coordination game, in so doing there is a danger of mistaking acquiescence to fundamentally emergent phenomena with deeper forms of endorsement. Emergent "societal conclusions" are not a rationally-motivated consensus. We conclude that systems approaches also fail to scale up deliberative democracy.

### **5. Tocqueville's Worry: the Deliberative Virtues and the Perils of Centralisation**

A summary of the argument thus far is in order. The first section outlined the scale problem of deliberative democracy. The necessary conditions for deliberative will-formation can be met only in small face-to-face groups. Widespread adoption of deliberative practices will lead at best to the creation of a mass of mini-publics each of which might internally form or come close to forming a rationally-motivated consensus. It will not lead to any society-wide deliberative consensus. Sections 2-4 considered and critiqued various attempts of deliberative theorists to overcome the scale problem. This section takes a different tack. If the chapter thus far has highlighted the limits to deliberative democracy, the aim of this section is to say something about the conditions under which it might be possible. One does not have to see it as a regulative ideal for all people in all places to see the normative attraction of participatory democracy in social life. On the Tocquevillian view advanced in this section, contemporary deliberative theorists lack an adequate account of the conditions under which agents are motivated to engage in democratic deliberation. The view advanced here is that agents are normally motivated to create deliberative associations in order to solve prosaic and quotidian local problems. Extensive state activity, however, undermines the propensity for deliberative association insofar as it relieves communities of the burden of resolving these problems themselves. Thus far in this chapter we have scoped out motivational concerns and taken for granted that citizens would want to engage in active deliberation. In this section we relax that assumption and try to identify the conditions under which citizens would have both sufficient knowledge *and* motivation to deliberate. Our

---

<sup>458</sup> Hence many might grudgingly follow anti-discrimination norms in their workplaces but elsewhere complain about excessive "political correctness".

contention is that deliberative theorists do not provide a satisfactory account on either of these two fronts.

They are typically concerned with deliberation about the major political topics of the day. Consider the following list of deliberative polls conducted by Fishkin:

Nine of these have been national – five in Britain (on crime, Britain's future in Europe, the monarchy, the 1997 British general election, and the future of the National Health Service on its 50th anniversary in 1998), one in Denmark (on the Euro right before the referendum in 2000), two in Australia (one before the republic referendum in 1999 and one on reconciliation with the aboriginals in 2001).<sup>459</sup>

He is not the only one. Deliberative democrats are overwhelmingly focused on agents debating in their role as *citizens* the merits of different proposals for state legislation. However, a cursory knowledge of the literature on voter ignorance and apathy, already mentioned in chapter three, should engender scepticism about whether most citizens would have sufficient motivation to engage in regular political deliberation. If many are already disincentivised to vote or to conduct basic research to inform their decision when they do vote, it is hard to imagine them giving up too many evenings to political deliberation.<sup>460</sup>

The problem though is not simply that voters might lack the motivation to discuss these grand issues of state legislation. There is an epistemic issue too. Determining, say, whether or not a country should join a currency arrangement or the future of a state-run health system or how to combat crime are very complex and difficult matters. The scale and complexity of such issues mean that even the putative experts struggle to understand them properly or diagnose the correct solutions. Non-experts stand even less of a chance of debating these matters in an informed manner. By contrast, let us imagine the residents of Swiss village deliberating on an amendment to the rules governing their communally owned land and roads.<sup>461</sup> They certainly have a rational choice incentive to do so, since the rules will affect resources that are important for their livelihood. They also are epistemically well placed to do so on the basis of their extensive local knowledge.<sup>462</sup> Rules governing communally owned land are certainly a more prosaic and “day-to-day” topic of deliberation than a discussion of how to best instantiate the difference principle. However, it is precisely

---

<sup>459</sup> (Fishkin 2003, 129)

<sup>460</sup> It is telling the degree to which Fishkin is a deliberative theorist who effectively has to *contrive* opportunities for political deliberation, be these through experimental deliberative polls or the more extensive deliberation day.

<sup>461</sup> The Swiss village is based, of course, on a case study of Elinor Ostrom (1990, 61–65). It is one of many examples she has found of communities working together to solve common problems.

<sup>462</sup> Local knowledge that is in some cases passed on through the generations.

these more prosaic issues about which real citizens will likely be motivated and informed enough to associate over and deliberate. If we wish to inculcate the deliberative virtues, our main concern should be giving citizens more control over such day-to-day matters.

This, I believe, is the great lesson of Tocqueville's *Democracy in America*, to which we will now turn. If the propensity to deliberate democratically is less pronounced than one might like, Tocqueville teaches us that this could stem from the scale and scope of contemporary state activity. He shares with contemporary deliberative democrats the general view that associational participation has the salutary effect of extending individual objectives beyond their own immediate interest to a wider sense of the common good.<sup>463</sup> He believes that such participation is necessary for a healthy society and to prevent social atomism. He thus distinguishes between *individualism* and *self-interest properly understood*. His attitude toward the former is cautious. Individualism is the

[Calm] and considered feeling which persuades each citizen to cut himself off from his fellows and to withdraw into the circle of his family and friends in such a way that he thus creates a small group of his own and willingly abandons society at large to its own devices.<sup>464</sup>

It is distinct from the vice of *egoism* – “an ardent and excessive love of oneself” – but acts as, one might say, a gateway to it.<sup>465</sup> On Tocqueville's score, an individual learns their self-interest properly understood when the distinction they make between their own welfare and that of others becomes porous. They would then sacrifice for others not on the basis of pure benevolence but under the awareness “that such sacrifices are as necessary to him who imposes them upon himself as to him for whose sake they are made”.<sup>466</sup> Like deliberative democrats, Tocqueville's inquiry is into the conditions under which citizens develop a genuine motivation to associate with their peers to further the common good. His answer, however, was *local* associational life. He greatly admired both the sheer quantity and variety of the associations he found in the 19<sup>th</sup> century United States, writing of the Americans that “not only do they have commercial and industrial associations to which all belong but also a thousand other kinds, religious, moral, serious, futile, very general and very specialised, large and small”.<sup>467</sup>

---

<sup>463</sup> Indeed, he is often an influence on more “communitarian” deliberative democrats e.g. (Sandel 1998).

<sup>464</sup> (Tocqueville 2003 [1840], 587)

<sup>465</sup> (Tocqueville 2003 [1840], 588)

<sup>466</sup> See the chapter entitled “How the Americans Combat Individualism by the Principle of Self-Interest Rightly Understood”. (Tocqueville 2003 [1840], 609–13)

<sup>467</sup> (Tocqueville 2003 [1840], 596)

Associational participation teaches citizens their self-interest properly understood. For a commitment to the general welfare does not come automatically but has to be nurtured. The nurturing process, however, must be grounded in the immediate interests of citizens: “[you] first get involved in the general interest by necessity, and then by choice; what was calculation becomes instinct; and by working for the good of your fellow citizens, you finally acquire the habit and taste of serving them”.<sup>468</sup> In other words, one is motivated to collaborate with others initially for self-interested reasons which only later develops into an intrinsic commitment to the welfare of others. Associational life also teaches citizens how to work together and solve problems. Robert Putnam famously transmuted this Tocquevillian notion into the parlance of contemporary social science with the term “social capital”.<sup>469</sup> Elinor Ostrom, already cited above, documented many examples of communities collaborating in a Tocquevillian manner to provide themselves with communal goods.<sup>470</sup>

Conventional political assumptions associate doctrines of limited government with a self-interested individualism and an activist state with the expression of other-regarding communal values. Tocqueville turns this dichotomy on its head. He thinks an activist state undermines associational life and promotes an undesirable form of atomistic individualism. For Tocqueville, necessity is the mother of both invention and sociability. Local association loses its *raison d'être* once state bureaucracies displace its role.<sup>471</sup> Quite simply, there is no need for individuals to solve a problem through free association if the state will do it instead. Once this incentive to collaborate is diminished, individuals are less likely to develop their self-interest properly understood and communities less likely their social capital. One might object that the central state provides its own avenue for public participation and collaboration. The displacement of more localised associations merely replaces one avenue for meaningful collaboration with another. However, on the view of Tocqueville and this thesis, such an objection fails because of an asymmetry between collaboration via the state and collaboration via local associations. As this chapter and thesis have argued, democratic citizens lack sufficient mutual knowledge to meaningfully collaborate with one another. By contrast, members of local associations often do have sufficient such knowledge.

Tocqueville saw a particular danger when this asymmetry is overlooked and the state wrongly seen as a vehicle for the common good. Citizens would be liable to a particular form

---

<sup>468</sup> (Tocqueville 2003 [1840], 893)

<sup>469</sup> (Putnam 2001)

<sup>470</sup> (Ostrom 1990)

<sup>471</sup> For a 20<sup>th</sup> century articulation of this Tocquevillian concern, see Nisbet (1953).

of “soft despotism”. The *despotism* of the contemporary state stems from the sheer scope of the activities undertaken by it. It claims the right “to a monopoly of all the force within the community, to make war, to make peace, to conscript life, to tax, to establish and dis-establish property, to define crime, to punish disobedience, to control education, to supervise the family, to regulate personal habits, and to censor opinions”.<sup>472</sup> As the state takes over so many swathes of life, the scope for autonomous democratic action decreases. Hence his famous lament:

It provides their security, anticipates and guarantees their needs, supplies their pleasures, directs their principal concerns, manages their industry, regulates their estates, divides their inheritances. Why can it not remove from them entirely the bother of thinking and the troubles of life?<sup>473</sup>

The despotism would be of a *soft* nature insofar as *democratic* citizens believe that the exercise of state power is nothing more than the expression of their own will. Indeed, insofar as the expansion of state power involves the application of uniform laws that appear to constrain all equally, citizens are likely to acquiesce and even promote it since “each citizen thinks that he is depriving all his fellow men of these powers that he grants to the crown”.<sup>474</sup> In a passage that seemingly anticipates the later thought of Michel Foucault, he writes:

I have always believed that this type of organised, gentle, and peaceful enslavement just described could link up more easily than imagined with some external forms of freedom and that it would not be impossible for it to take hold in the very shadow of the sovereignty of the people.<sup>475</sup>

Deliberative theorists today, by and large, have lost sight of these Tocquevillian concerns. They tend to favour welfare states that are more active and extensive than the status quo. The oft stated concern is that deliberative equality would not be possible save for radical state intervention to guarantee much greater socio-economic equality.<sup>476</sup> The overlooking of

---

<sup>472</sup> The citation belongs to Walter Lippmann. He is cited by Cavanaugh (2011, 28). This section discusses the state as if it were a singular entity as a shorthand. This is not to overlook, of course, the point made in chapter two that the state is better understood as a network rather than a unitary organisation.

<sup>473</sup> (Tocqueville 2003 [1840], 805–6)

<sup>474</sup> (Tocqueville 2003 [1840], 782)

<sup>475</sup> (Tocqueville 2003 [1840], 806)

<sup>476</sup> A clear statement of this view is offered by Bohman (1997). As David Schmidtz (1997) observes, it is often thought that state intervention is needed to “guarantee” a desired outcome whose attainment would otherwise be precarious. Thus, a deliberative democrat like Bohman thinks that we need the state to guarantee socio-economic equality, without which deliberative practices will not flourish. Schmidtz, however, observes that the “politics of guarantees” (my term) often has unintended consequences. One is that the interventions might, in fact, fail to provide guarantees. Another is that they strip citizens of the responsibility to try and solve problems themselves. It might be that citizens are motivated to deliberate only when there is a felt need. If this is right, attempted state guarantees of socio-economic equality might be in tension and not complementary with the aim

these Tocquevillian concerns is unfortunate. There is evidence to support the displacement thesis, suggesting that deliberative theorists have overlooked the costs of the rise of welfare bureaucracies.

A good historical example of the displacement with which Tocqueville is concerned is the decline of friendly societies in the United States and the United Kingdom following the expansion of state welfare in the twentieth century.<sup>477</sup> Friendly societies were voluntary associations formed by men and women for a whole range of welfare purposes, from insurance in cases of infirmity and sickness to the care of the elderly to providing funding for burials. They also provided medical services. They were distinct from charities since all members paid contributions to the society and could claim funds when the need arose as a matter of entitlement. They emerged and developed in the nineteenth century. And they were popular. The 1911 National Insurance Act mandated social insurance in the United Kingdom. However, David Green estimates that at least 9 of the 12 million people covered by the Act were already members of friendly societies.<sup>478</sup> These societies weren't simply insurance schemes. These were self-governing democratic associations created by their predominantly working class members.

The constitution of the typical friendly society treated all members as equal under the rules of the society and all were expected to contribute to its governance.<sup>479</sup> Indeed, both the day-to-day governance and the rules of the association themselves were determined democratically by their members.<sup>480</sup> New members particularly were encouraged to take part in these democratic processes and not view their participation solely instrumentally as means to secure insurance.<sup>481</sup> A local branch would have various elected positions, ranging from warden to financial secretary. The member of a typical branch was eligible and indeed expected to seek to occupy these positions. Green describes them as "training grounds for democracy," particularly so for manual workers whose day jobs may have been mundane.<sup>482</sup> These were associations in which deliberative ideals were met and were done so on a mass

---

of promoting deliberative practices. The example of the friendly societies, wherein working men and women organised in deliberative associations to provide themselves with welfare in the *absence* of guaranteed state provision, analysed further below, will provide some substantiation for this claim.

<sup>477</sup> (Green 1993); (Beito 2000); (Davies 1997). What follows draws primarily upon Green's account of the rise and fall of UK friendly societies.

<sup>478</sup> (Green 1993, 26)

<sup>479</sup> (Green 1993, 31)

<sup>480</sup> (Green 1993, 34)

<sup>481</sup> (Green 1993, 30)

<sup>482</sup> (Green 1993, 33)

scale. The workers had a strong material incentive to participate insofar as their livelihoods were at stake. They were also well placed epistemically to deliberate about these matters. They would know the risks facing persons in their social class and occupation. This would inform their deliberation and decision-making as to the appropriate rules for their association. The important point here is that such deliberations were embedded in concrete social practices. The objects of discussion were not abstract matters removed from day-to-day concerns. Relatedly, the motivation to join these associations, at least initially, did not come from an abstract commitment to civic virtue. It came from a concrete concern to protect themselves and their families from misfortune and hardship. This provides substantiation for Tocqueville's view that the motivation to engage in deliberative practices is best built on top of such "mundane" and "prosaic" concerns.<sup>483</sup>

The deliberative democrat might be tempted to dismiss the friendly societies as insufficiently political. These were not citizens deliberating about major matters of state. Indeed, many of their nineteenth century members would not have had the right to vote in political elections. However, this response would beg the question as to what counts as "political". The welfare functions performed by the friendly societies are a large if not the main bulk of state activities today. Though the friendly societies did not directly influence state legislation, they fulfilled many functions that we now think of as political. Arguably they gave their members much more control over how these services were received than do contemporary state welfare institutions.<sup>484</sup> Green locates the decline of friendly societies in rent seeking by the medical profession and for-profit insurance companies, who lobbied so that the 1911 legislation would favour their prospects over those of the societies.<sup>485</sup> However, and consistent with Tocqueville's displacement thesis, the subsequent rise of the welfare state is surely also an important factor. There is no point in self-organising to provide oneself with unemployment insurance or healthcare if these are guaranteed by state provision. Thus, institutions that enabled and motivated agents to take part in deliberative self-governance on a mass scale were likely undermined by expansive state activity.<sup>486</sup>

---

<sup>483</sup> The scare quotes are inserted because these concerns are not mundane for those who are first-personally affected by them. They are mundane from the perspective of the political philosopher who wants to ask what the just society looks like from some Archimedean perspective.

<sup>484</sup> Many friendly societies provided their members with medical services by contracting directly with doctors. The British Medical Association complained that the societies "underpaid" their doctors, suggesting their bargaining power was significant. (Green 1993, 62)

<sup>485</sup> (Green 1993, chap. 9)

<sup>486</sup> One may believe, as I do, that some of the shift to a more centralised welfare system was justified on an all-things-considered evaluation. But the point here is that a diminishment of participatory democracy and associational life are to be considered significant costs in that evaluation.



## Conclusion

This chapter has argued that deliberative democracy's scale problem cannot be solved. Consequently, a large-scale deliberative demos is not a feasible prospect. The deliberative ideal is best met in a decentralised social order characterised by comparatively restrained central state intervention. So concludes the second part of this thesis and its epistemic critique of descriptive and aspirational associationalism. This and the previous chapter have emphasised that it is mutual ignorance rather than a lack of motivation to engage in politics, or a lack of consensus on the right principles, that inhibits collective political agency. Per the emergent view defended in this thesis, the latter is the prime cause of alienation among liberal citizens. In the final part of this thesis we will consider the implications of this argument for the liberal self. Collective political agency is attractive in part because it promises social unity to the individual. She can see her life-plan as at least partly a contribution to a grand project she shares with all her fellow citizens. The potential anomie of an impersonal liberal order is thereby alleviated. The thought that political participation contextualises individual achievements as a contribution to a greater whole dates back to Hegel. Rawls gives this idea a contemporary construal when claiming that the well-ordered society is a "social union" whose citizens participate in one another's achievements. Focusing on Rawls' account, chapter six will argue that this desire for social unity cannot be met. The liberal self cannot see her own life as a contribution to a greater project. She is too ignorant of the circumstances of her fellow citizens. Epistemic alienation is the price we pay for living in large-scale orders. In chapter 7 we turn to alleviate Hayek's worry that liberal agents will be unable to accept a liberal order because they consider this price too high to pay and that the stability of a liberal order will be cast into doubt as a consequence.

### **Part III: Alienation and the Liberal Self**

## Chapter Six: Social Unity and the Well-Ordered Society

### Abstract

This chapter considers the claim that membership of political association can reconcile subjective freedom with community. This claim, dating back to Hegel, manifests itself in various ways in contemporary political theory. Rawls gives it the most elaborate and explicit formulation. He characterises a well-ordered society as a preeminent *social union*. Members of a social union gain recognition for their expression of human excellence and enjoy the expression of excellence by others. Rawls oscillates between two different conceptions of a social union and in turn two different explanations of why the well-ordered society is the preeminent social union. While not denying that social union is possible in civil society associations, this chapter critiques both versions of Rawls' claim that the well-ordered society is the preeminent social union or indeed a plausible social union of any sort.

### Introduction

Weber's distinction between the communal and face-to-face *Gemeinschaft* and the impersonal and bureaucratic *Gesellschaft* is well established in Western thought. We can read Hayek as arguing that socialism and social justice constitute the erroneous application to the *Gesellschaft* of the ethics the *Gemeinschaft*. One way of understanding the shift from small-scale community to impersonal relations is in terms of a trade-off of communal bonds for individual freedom. Hegel's *Elements of the Philosophy of Right* is predicated on this claim.<sup>487</sup> For Hegel, Plato's *Republic* exemplified a communal social order with little room for individuality. Each person occupies a fixed and ascriptive social role. The function of each social role – the producers, the auxiliaries, and the guardians – harmonise with one another. There is no notion of a private interest distinct from one's socially recognised function within the community. Hegel compares-and-contrasts this social world of the *Republic* with the impersonality of commercial society. Social coordination there is facilitated by contract and exchange between individuals pursuing private aims. The situation here is the reverse of Greek life for Hegel. Subjective individuality flourishes but seemingly at the expense of the social harmony found in the communal grouping.

Assuming there is no going back to the old social order, even if that was thought desirable, the questions remaining are whether we must make do with this loss of community or whether we can reconcile community with the *Gesellschaft*. Hegel famously thought the latter was possible and indeed necessary. While his favourable attitude towards individual

---

<sup>487</sup> (Hegel 2011 [1896])

or – in his parlance – “subjective” freedom is not to be doubted, he was concerned about the alienation it caused. The issue for Hegel is that individuals do not simply want to express their idiosyncratic individuality. They wish to express that freedom in a manner that harmonises with the equivalent expression of their peers. This is the crucial good provided by one’s membership of a state. Hegel here distinguishes between the institutional manifestations of the state, such as the judiciary or the civil service – “the political state proper” – from the state as a manifestation of society as one integrated ethical community – *Sittlichkeit*. Political institutions should be rationally designed and structured such that otherwise disaggregated individual decision-making coheres into a purposeful and meaningful harmony. In his own constitutional proposals Hegel espouses the creation of guild-like corporations. On this corporatist model, our individual freedom to join any kind of guild is assured but the interactions between them are subject to centralised coordination to ensure that each guild contributes to the universal good. The end result is that the exercise of subjective freedom by each harmonises with the universal good of all. The individual experiences not only subjective freedom but participation in this wider and integrated *Sittlichkeit*. Individuality *and* community are secured.

Now, the precise details of Hegel’s proposal need not detain us. It is the underlying idea that matters. The Hegelian claim that political membership is a surrogate for community in an impersonal social order remains influential. Recall the claim by Margaret Moore that people who “exercise collective self-government” thereby “shape the conditions of their existence, and their future together”. They achieve a “collective” dimension of autonomy not available to those whose freedom is restricted solely to the “individual private sphere”.<sup>488</sup> Yael Tamir argues that membership of a nation-state can elevate one’s personal achievements. Assuming that one has a strong national identity, then their writing a good novel is not merely an isolated act of creativity. It is a contribution to, say, *American* literature. Indeed, in popular parlance we find nationalist references to “our” playwrights or singers or athletes and so on. In Tamir’s words, such national identification:

[Contextualises] human actions, no matter how mundane, making them part of a continuous creative effort whereby culture is made and remade. Insofar as every action within a national entity is endowed with this additional dimension, individuals living within national frameworks enjoy options unavailable to those outside them.<sup>489</sup>

---

<sup>488</sup> (Moore 2015, 66)

<sup>489</sup> (Tamir 1993, 85)

In a similar vein, various political philosophers have argued that the effective exercise of autonomy requires membership of a wider “national” or “societal” culture.<sup>490</sup> John Rawls in particular has tried to update this Hegelian idea with his characterisation of a well-ordered liberal society as a social union. He invokes this idea to show that the well-ordered society can account for the value of community. The basic idea is that we can arrange liberal institutions so that each citizen feels that their exercise of liberty coheres into a greater unity. Their private plans being experienced as “a plan within a plan, this superordinate plan being realized in the public institutions of society”.<sup>491</sup> The hope of reconciliation between individuality and community comes naturally to the associational view. If citizens are indeed capable of acting together in the pursuit of shared ends, then each can contextualise their personal actions as a contribution to this wider endeavour. On the emergent view, however, the mutual ignorance of our respective goals and circumstances renders this too much to hope for. One can, of course, find community in the epistemically tractable context of civil society associations but not within the context of impersonal cooperation with millions of fellow strangers. This chapter offers an immanent critique of Rawls’ characterisation of the well-ordered society as a social union.

As we shall see, Rawls’ discussion is not without its exegetical difficulties. It is articulated in the third part of *A Theory of Justice*, a section of the book in which Rawls displays an “uncharacteristic lack of clarity” according to Samuel Freeman.<sup>492</sup> Accordingly, any reading of the social union argument will have to engage in some exegesis to make up for the ambiguities in Rawls’ own presentation. This chapter argues that Rawls oscillates between two distinct models of what a social union is, though he does not render the distinction between them explicit. This has to be reconstructed from his various remarks on the matter. Depending on the passage-in-question, he draws on both models when claiming the well-ordered society is a social union. This chapter will argue that on neither of these two interpretations can the well-ordered society be understood as such. The textual ambiguities notwithstanding, Rawls arguments here deserve consideration because they are an elaborate if not always fully clear attempt at showing how liberal individuality and large-scale community might be reconciled. The argument to follow is divided into four sections. Section one outlines in general terms Rawls’ understanding of a social union and the conception of the good on which it is based. Section two outlines the two interpretations of a social union

---

<sup>490</sup> (Kymlicka 1996, chap. 5); (Margalit and Raz 1990)

<sup>491</sup> (Rawls 1999, 463)

<sup>492</sup> (Freeman 2007, 277); the bulk of the analysis to follow will draw on Rawls’ discussion of social unions in that last third of *Theory*.

to be found in *Theory*. These are the *completeness* and *emulation* interpretations. Section three critiques the claim that the well-ordered society can be understood as an emulation social union. Section four critiques the claim that it can be understood as a completeness social union. The chapter concludes by contrasting Rawls' characterisation of a liberal order as a social union with Hayek's markedly different vision of a liberal order. The argument to follow takes for granted Rawls' claim that the desire for social union is widespread and that such unions can be found in civil society. It questions whether the well-ordered society as a whole can be modelled as such on the basis of the epistemic concerns of this thesis.

### 1. Goodness as Rationality

Rawls' understanding of a social union is based on his conception of "goodness as rationality". This is the same broad but thin account of the good life from which his theory of primary goods is derived. If, for Rawls, primary goods are goods one wants whatever else one wants, goodness as rationality is the structure one's conception of the good would take whatever else is its content. It is outlined in chapter 7, part III of *Theory*. It assumes that a worthy life is a planned one. Following the idealist philosopher James Royce, Rawls assumes that a "person may be regarded as a human life lived according to a plan".<sup>493</sup> A rational life plan would be consistent with two formal criteria. It would be consistent with basic principles of rational choice and it would be selected "with full awareness of the relevant facts and after a careful consideration of the consequences".<sup>494</sup> Though these two formal criteria do not tell us anything about the kind of plans rational agents might choose. Rawls tries to meet this desideratum with the following claim about human motivation. Individuals tend to pick those life plans that accord with the *Aristotelian principle*:

Other things equal, human beings enjoy the exercise of their realized capacities (their innate or trained abilities), and this enjoyment increases the more the capacity is realized, or the greater its complexity. The intuitive idea here is that human beings take more pleasure in doing something as they become more proficient at it, and of two activities they do equally well, they prefer the one calling on a larger repertoire of more intricate and subtle discriminations.<sup>495</sup>

---

<sup>493</sup> (Rawls 1999, 358); there is little room for serendipity and surprise in this theory of the good life. For an excellent critique of Rawls on this score, see Larmore (1999).

<sup>494</sup> (Rawls 1999, 358–59)

<sup>495</sup> (Rawls 1999, 373)

According to the principle, then, chess would be preferred to checkers and algebra to basic arithmetic.<sup>496</sup> The proof of the principle is allegedly to be found in common sense observation. Rawls notes that it “seems to be borne out by the many facts of everyday life, and by the behaviour of children and some animals”.<sup>497</sup> Rawls does not overlook the obvious objection that many appear to lead lives that are not in accordance with the principle. He is careful to formulate it as a “tendency and not an invariable pattern of choice” and one that can be over-ridden by “countervailing inclinations”.<sup>498</sup> As a principle of motivation, it seems quite reminiscent of Mill’s distinction between the higher and lower pleasures.<sup>499</sup>

Now, if there is an objection to be made against the principle, it is that it seems to be an excessively inward-looking account of motivation. One can readily point to examples of plausible and reasonable life-plans that place at best instrumental value on the development of one’s talents. A nurse’s principal motivation, surely, is not to develop their nursing capacities to ever higher forms of complexity and nuance for their own sake but, quite simply, to help the sick recover their health. A scientist might well take intrinsic pleasure from improving their capacities of observation, testing, and so forth. But suppose their primary goal is to help cure cancer, in which case they might prefer to work in that field even if a less socially pressing field of scientific inquiry would allow them to express or develop their scientific skills to a higher degree of complexity. Many opt for a mundane job that does not optimise their capacities but does help them to meet other goals such as supporting their family. Indeed, we can think of occupations or activities in which it would seem socially irresponsible for one to be motivated by the Aristotelian principle. Imagine a surgeon who, all other things equal, preferred to perform technically challenging but medically non-essential operations over mundane but life-saving ones. The principle taken on its own, then, appears to conjure an image of individualists who take pleasure only in the realisation of their own capacities but place little weight on social bonds with others.

Rawls is aware of this potential objection and attempts to avoid it at the outset by pairing the principle with an account of community. Our self-respect, he writes, is dependent on having a rational plan of life in accordance with the Aristotelian principle *and* “finding our person and deeds appreciated and confirmed by others who are likewise esteemed and their

---

<sup>496</sup> (Rawls 1999, 375)

<sup>497</sup> (Rawls 1999, 376)

<sup>498</sup> (Rawls 1999, 376)

<sup>499</sup> Indeed, other than Aristotle himself, Rawls cites Mill’s discussion of the distinction between lower and higher pleasures in chapter 2 of *Utilitarianism* as influencing his thinking here. He claims that in that discussion Mill “comes very close” to stating the Aristotelian principle. (1999, 374 fn 24)

association enjoyed”.<sup>500</sup> We don’t simply want to exercise our capacities to a high degree of complexity but be *socially recognised* for doing so. It is impossible to maintain one’s commitment that one’s rational-life plan is worthwhile unless one gains such recognition.<sup>501</sup> The development of only certain kinds of capacities engenders such recognition however. These capacities will not be what Rawls calls *exclusive* capacities that are good primarily for the capacity-bearer but *excellences* that are good both for the bearer and other members of the community.<sup>502</sup> What makes a capacity an excellence as opposed to a mere exclusive good? An answer can be gleaned from Rawls’ “companion affect” to the Aristotelian principle:

[As] we witness the exercise of well-trained abilities by others, these displays are enjoyed by us and arouse a desire that we should be able to do the same things ourselves. *We want to be like those persons who can exercise the abilities that we find latent in our nature.*<sup>503</sup>

I achieve an excellence when I manifest a capacity that is latent in other members of the community. In turn I enjoy seeing others exercise capacities latent in my own nature. Putting this all together, we can say that individuals motivated by the Aristotelian principle and the related need for social recognition will seek to develop those latent capacities that they share with others. They will associate for this purpose and Rawls, following Wilhelm von Humboldt, calls these associations “social unions”. These associations come in very many forms. The associations of “art and science, of religion and culture of all kinds, high and low” can be thought of as social unions, the “essential thing is that there be a shared final end and accepted ways of advancing it which allow for public recognition of the attainments of everyone”.<sup>504</sup> Rawls is at pains to distinguish the motivation of agents in a social union from that of agents in commercial exchange. Purportedly, agents in a social union are solely intrinsically motivated to realise the shared end of the union whereas agents in commercial exchange are solely extrinsically motivated to maximise monetary income.<sup>505</sup>

---

<sup>500</sup> (Rawls 1999, 386)

<sup>501</sup> (Rawls 1999, 387)

<sup>502</sup> (Rawls 1999, 387)

<sup>503</sup> (Rawls 1999, 375–76); emphasis added

<sup>504</sup> (Rawls 1999, 461)

<sup>505</sup> (Rawls 1999, 457–59)



## 2. Two Models of a Social Union

As already mentioned, Rawls' discussion of a social union is not quite as pellucid one might like. On this chapter's reading, there are two distinct definitions of what a social union is and by extension the manner in which the well-ordered society is a social union. The first is the completeness model. This model is the most prominent in Rawls' writing. Indeed, when formally defining a social union, Rawls opts for this model. The other is the emulation model. Both will now be outlined in turn. The completion model takes its inspiration from Wilhelm von Humboldt. Like Rawls, von Humboldt is a liberal individualist who values the development of individual capacities but also wishes to provide an account of how such individuality can be reconciled with community and sociability. In his landmark essay *The Sphere and Duties of Government* (sometimes translated as *The Limits of State Action*), von Humboldt observes that for want of time individuals can develop only a subset of the full range of their latent capacities. This is undesirable because it represents a "partial cultivation" of the self.<sup>506</sup>

Von Humboldt proposes two solutions to this problem. One is the harmonious development of a person's talents over a life time, such that different powers are developed at different stages. The other is to collaborate with individuals who have developed different talents, such that each individual can derive enjoyment from a greater array of talents than they could have developed on their own.<sup>507</sup> Rawls is interested in the second solution. He writes that "it is through social union founded upon the needs and potentialities of its members that each person can participate in the total sum of realized natural assets of the others".<sup>508</sup> As a "pure form" of a social union he gives the example of an orchestra, in which "each have by a kind of tacit agreement set out to perfect their skills on the one [instrument] they have chosen so as to realise the powers of all in their joint performances".<sup>509</sup> The core idea is that "persons need one another since it is only in active cooperation with others that one's

---

<sup>506</sup> (Humboldt 1854, 13). The partial cultivation of one's capacities is only really a problem if we presume the validity of something like the Aristotelian principle. If one thinks that cultivation of one's talents is intrinsically good, then talents left *uncultivated* might be a problem. They are a missed opportunity to realise goodness. It is also more likely that one will experience the peril of uncultivated talents if one also thinks that the cultivation of one's talents to a *high degree of sophistication* is intrinsically good. For leaving many talents uncultivated will probably be the opportunity cost of developing some talents to a high degree of sophistication. (Someone who thinks that cultivating their talents but only to a *moderate* degree is intrinsically good will less likely have the problem of partial cultivation. Since they might have more time to develop the full range of their talents to a satisfactory degree.)

<sup>507</sup> (Humboldt 1854, 13)

<sup>508</sup> (Rawls 1999, 459)

<sup>509</sup> (Rawls 1999, 459n4)

powers reach fruition. Only in a social union is the individual *complete*".<sup>510</sup> Hence, our terming it the completeness model. The following scenario captures, I believe, what Rawls has in mind here:

Rosie and Jim are members of a world class heavy metal band. Each had the talent to either be an excellent drummer or bassist. As it happens, Jim dedicates himself to drumming and Rosie to bass guitar. They take great pleasure in playing together. When Jim performs with Rosie, he feels that he is actively expressing his latent talent for playing the bass guitar as well as his manifest talent for drumming. The vice-versa applies to Rosie. It is as if each completes the other by allowing them to express through the jointly-executed performance those latent talents they forewent developing as the opportunity cost of those they did develop and manifest.

On the completion model, members of the social union manifest *different* capacities that are latent in all of them. Social union is experienced as good because it completes the self by realising one's otherwise uncultivated talents. Though we should note an important necessary condition. When Rawls says that "persons need one another since it is only in *active cooperation* with others that one's powers reach fruition," I take him to mean that one participates in the expression of one's latent talents by others *iff* they share causal responsibility for its expression. In other words, Jim would not be participating in Rosie's expression of her actual (and his latent) talent for bass playing if he was just a *spectator* in the crowd and Rosie was playing with a different drummer. Jim participates in her playing the bass guitar insofar as they play a song together. Or to return to Rawls' own example, it is insufficient to just watch the orchestra play. One has to play in it. We can also express this point using Bratman's notion of a joint-intentional action. Rosie and Jim share the intention that they play "Master of Puppets" *iff* each knows that the other intends that they play the song and intends to play their part in playing the song and that this is common knowledge between them. Under such a scenario, each would know that they are part-authors of the contribution of the other. It is under such conditions where the phenomenological experience of participating in one another's talents seems to take place. We shall return to this all-important point below.

Rawls has other passages in which his description of the social union is inconsistent with the completion model. Recall that in the aforementioned citation describing the Aristotelian principle's "companion affect," Rawls writes that "as we witness the exercise of well-trained abilities by others, these displays are enjoyed by us and arouse a desire that we should be able to do the *same things* ourselves".<sup>511</sup> Members of an orchestra do not wish to do the

---

<sup>510</sup> (Rawls 1999, 460 n4); emphasis added.

<sup>511</sup> (Rawls 1999, 375); emphasis added.

*same thing*. They play different instruments after all. Similarly, he writes that competitive games can also be seen as social unions since the “public desire to execute a good and fair play of the game must be regulative and effective if everyone’s zest and pleasure are not to languish”.<sup>512</sup> But in many competitive games, such as chess or boxing, the agents develop and cultivate broadly the same set of capacities rather than different ones. If they are social unions, then, they do not appear to be of the same sort as an orchestra or a band. For this reason, we can read Rawls as invoking a different understanding of a social union in these passages, the *emulation* model. The following scenario captures, I believe, what Rawls has in mind here:

Alf and Bettie are professional chess players. They compete in a national chess tournament. Bettie beats Alf by making a surprise and highly creative series of moves. Alf, however, is magnanimous in defeat. He is impressed by Bettie’s display of excellence in the activity to which they are both dedicated. He feels that playing against someone of her capability makes him a better chess player. Bettie, for her part, is grateful for having a challenging opponent in Alf, without whom she would not have exhibited the creative play that she did.

On the emulation model, the parties manifest *broadly the same capacities*. Social union is experienced as good because seeing others excel in an activity to which one is dedicated is joyful – even perhaps when they are competing against you – and potentially provides a model for emulation. A good practical example might be Roger Bannister’s achievement of being the first to run a sub-four-minute mile, soon after which many followed in his wake. Both of these models of social union serve the same purpose of showing why individualists dedicated to cultivating their own capacities would have strong reasons to enter a social union. On the completion account, social union allows me to participate in the expression of those capacities that I forewent developing as the opportunity cost of developing my own. On the emulation account, social union allows me to further develop and refine those capacities I did choose to develop by emulating the exemplars.

This chapter accepts for the sake of argument Rawls’ account of goodness as rationality and the attendant claim that we desire social union.<sup>513</sup> Supposing all persons had conceptions of

---

<sup>512</sup> (Rawls 1999, 461)

<sup>513</sup> This is not to say it is without its problems. It exaggerates the importance of planning and overlooks the importance of serendipity to the good life. It is also implausibly monistic in assuming that the realisation of our inherent capacities is central if not primary to any reasonable conception of the good life. As I suggested in some counter-examples above, this just seems mistaken. A more plausible account of human motivation would treat as reasonable a wider range of motivations, including extrinsic ones. I personally believe that Deirdre McCloskey (2011) is a surer guide than Rawls here. Still, and in keeping with the methodology of this thesis, we imagine a society where all do subscribe to such a conception of the good to be able to isolate the epistemic rather than motivational barriers to social union.

the good consistent with goodness as rationality, the question then is whether the well-ordered society can plausibly be regarded as a social union. Rawls not only answers in the affirmative but also claims it to be the “social union of social unions” that “realizes to a *preeminent* degree the various forms of human activity”.<sup>514</sup> In some passages of *Theory* Rawls describes the well-ordered society as a social union along the completion model and in others as one along the emulation model. There are, then, two accounts of the kind of social union the well-ordered society is and why it is preeminent. We shall outline both of these in the remainder of this section before subjecting them to critique in the following section.

The well-ordered society is described as a completion social union in §86, “The Good of the Sense of Justice”. Rawls writes:

Because such a society [i.e. The WOS] is a social union of social unions, it realizes to a preeminent degree the *various* forms of human activity; *and given the social nature of mankind, the fact that our potentialities and inclinations surpass what can be expressed in any one life, we depend upon the cooperative endeavours of others not only for the means of well-being but to bring to fruition our latent powers.* And with a certain success all round, each enjoys the *greater richness and diversity of the collective activity.*<sup>515</sup>

Similarly in §85, “The Unity of The Self”, he writes

... [A]s shown by the notion of society as a social union of social unions, the members of a community participate in one another’s nature: *we appreciate what others do as things we might have done but which they do for us, and what we do is similarly done for them.*<sup>516</sup>

The well-ordered society is in these passages clearly described as a completion social union. Participation in it is good because it allows us to participate in the expression of our otherwise uncultivated talents. In *Political Liberalism* Rawls also speaks of a social union in this manner.<sup>517</sup> One might ask how Rawls squares this view with liberal neutrality. Rawls addresses this question in *Political Liberalism* by suggesting that one’s conception of the good will be enlarged by participation in the well-ordered society, provided it is consistent with justice and within a “certain wide range” of conceptions.<sup>518</sup> Why might the well-ordered

---

<sup>514</sup> (Rawls 1999, 500); emphasis added.

<sup>515</sup> (Rawls 1999, 500); emphasis added.

<sup>516</sup> (Rawls 1999, 495); emphasis added.

<sup>517</sup> (Rawls 1996, 320–23)

<sup>518</sup> (Rawls 1996, 323); the above qualification of a “wide range” of conceptions suggests an equivocation on Rawls’ part. We can easily imagine persons who will not benefit from the “social union of social unions” but whose conceptions of the good are otherwise “reasonable” from a Rawlsian perspective. Say persons who are willing to abide by liberal justice but who have a hermitic conception of the good or who regard the ways of life of mainstream society as depraved. It is hard to see persons

society be the *preeminent* completion social union? The clearest answer is to be found also in *Political liberalism*. There he argues that participation in the well-ordered society allows each to take part in a more “comprehensive good than the determinate good of individuals when left to their own devices or limited to smaller associations. Participation in this more comprehensive good can greatly enlarge and sustain each person’s good”.<sup>519</sup> Participation in, say, an orchestra allows me to participate in the expression of my uncultivated talents with respect to only one human activity in classical music. Participation in the well-ordered society, however, allows me to participate in the expression of my uncultivated talents in any number of altogether different disciplines, such as sport or art or philosophy.

Other passages describe the well-ordered society as an emulation social union. These are to be found in §79 “The Idea of a Social Union” and Rawls’ preliminary discussion of social unions in §67 “Self-respect, Excellences and Shame”. In §79 Rawls writes that in the well-ordered society “the successful carrying out of just institutions is the shared final end of all the members of society”.<sup>520</sup> He had earlier written in §67 that acting justly is intrinsically valuable per the Aristotelian principle. Such moral virtues are “excellences” in that they are “good from the standpoint of ourselves as well as from that of others”.<sup>521</sup> This discussion is referred back to in §79 where Rawls writes:

Now it is clear that these excellences are displayed in the public life of a well-ordered society. Therefore *the companion principle to the Aristotelian principle implies that men appreciate and enjoy these attributes in one another as they are manifested in cooperating to affirm just institutions*. It follows that the collective activity of justice is the preeminent form of human flourishing.<sup>522</sup>

The well-ordered society is here described as an emulation social union insofar as each citizen expresses the same capacity, i.e. for just conduct. Each takes pleasure from seeing the just conduct of the others and the just conduct of each is potentially a source of emulation for the others.

Why might the well-ordered society be the *preeminent* completion social union? Even if we hold that (a) justice is the relevant capacity to be realised and (b) liberal citizens want to express their capacity for justice and see others doing the same, participation in civil society associations or even being just in one’s day-to-day conduct seems sufficient. The need for

---

of either type having their conception of the good “enlarged” by participation in the well-ordered society. But I set this objection to one side.

<sup>519</sup> (Rawls 1996, 320)

<sup>520</sup> (Rawls 1999, 462)

<sup>521</sup> (Rawls 1999, 390)

<sup>522</sup> (Rawls 1999, 463); emphasis added.

the well-ordered society is unclear. Rawls' response here draws on his Kantian premises. For Rawls, just conduct is a good of excellence per the Aristotelian principle *insofar as it express our nature as free and equal moral beings*.<sup>523</sup> In §79 he writes: "Human beings have a desire to express their nature as free and equal moral persons, and this they do *most adequately* by acting from the principles that they would acknowledge in the original position".<sup>524</sup> Why does Rawls think this? Earlier in *Theory* he writes:

The description of the Original Position resembles the point of view of *noumenal selves*, of what it means to be a free and equal rational being. Our nature as such beings is displayed when we act from the principles we would choose when this nature is reflected in the conditions determining the choice.<sup>525</sup>

Since those principles govern the well-ordered society, we can only act from those principles and by extension express our free and equal nature by participating in it. A formal reconstruction of Rawls' somewhat obscure position might run as follows:

1. Persons wish to express their nature as a free and equal beings.
2. Persons express that nature when acting on the basis of an autonomously affirmed principle.
3. Persons act on an autonomously affirmed principle when the principle they act upon is justifiable to all persons understood as free and equal.
4. Persons act upon principles justifiable to all understood as free and equal *iff* they act upon principles derived from the original position.<sup>526</sup>

∴

5. Persons wish to express their nature as free and equal by acting upon principles derived from the original position.
6. One can act on the principles derived from the original position only by participating in the well-ordered society and complying with its regulative rules.<sup>527</sup>

∴

---

<sup>523</sup> (Rawls 1999, 390–91, 462)

<sup>524</sup> (Rawls 1999, 462–63); emphasis added.

<sup>525</sup> (Rawls 1999, 225); emphasis added. In *Political Liberalism*, Rawls (1996, 73) similarly writes that persons attain their "full autonomy. . . when they act from principles of justice that specify fair terms of cooperation they would give to themselves when fairly represented as free and equal persons".

<sup>526</sup> The veil of ignorance is designed to model the condition under which all are free and equal agents.

<sup>527</sup> Since the rules governing the basic structure of that society is the subject of the conception of justice derived from the original position.

7. Persons wish to express their nature as free and equal by participating in the well-ordered society.

The well-ordered society is the preeminent emulation union because can we express our nature as free and equal only by subscription to its rules of justice. Why might Rawls think this is not the case when we comply with rules of justice regulating civil society associations? Presumably because premise 3 is not met. In such settings one might be motivated to act justly because of fraternal links with one's peers rather than an abstract commitment to act on principles justifiable to all free and equal persons. Relatedly, the content of such principles often cannot be affirmed by all rational agents because they are based on partisan understandings of value. The principles of justice regulating a religious association, for instance, might be based on scriptural premises that not all agents can accept.

Note that these two conceptions of a social union are not necessarily mutually exclusive, but they are independent arguments. If one is found wanting, Rawls can turn to the other. We will treat them in turn. The basic argument in both cases is that our relationship as citizens is dubiously modelled as any kind of social union, let alone as the preeminent sort.

### 3. The Critique of the Emulation Interpretation

Even if we grant Rawls these contentious Kantian premises, it does not actually follow that participation in the well-ordered society is the preeminent emulation social union. Per Rawls' Kantianism, we wish to express our nature as free and equal beings by acting on principles that are justifiable to all. In Kant's terms, the relevant test here is whether the maxim one acts upon meets the categorical imperative. As Rawls himself notes however, Kant intended the imperative to apply to quotidian contexts of everyday life:

... Kant's account of the Categorical Imperative applies to the personal maxims of sincere and conscientious individuals in everyday life. To be sure, in the course of testing such maxims we are to compare social worlds, that is, the social world that results when everyone follows the proposed maxim, as if by a law of nature, with the social world in which the contradictory maxim is followed. But this comparison of social worlds is undertaken singly by each person and for the purpose of judging a given personal maxim.<sup>528</sup>

In other words, there are good Kantian reasons to think premise 4 from above is not correct. Acting upon principles justifiable to all as free and equal does not *necessarily* entail acting

---

<sup>528</sup> (Rawls 1980, 552)

upon principles that are derived from the original position and consequently intended to regulate the basic structure of a *political* society. One can act upon universally justifiable principles and so express their nature as a free and equal being by, say, simply keeping their promises. If so, being a citizen of a well-ordered society is not necessary to expressing one's nature as a free and equal being. A stateless person or an immigrant residing in a country where they lack citizenship could just as easily express their nature so understood.

There are also reasons to think that compared with civil society associations, the well-ordered society fares quite poorly as an emulation social union on Rawls' own terms. He writes that we admire and wish to emulate "moral exemplars" because they "display skills and abilities, and virtues of character and temperament, that attract our fancy". He goes on to say that this desire to emulate "is also a companion affect to the Aristotelian Principle, since we enjoy the display of more complex and subtle activities and these displays tend to elicit a desire in us to do these things ourselves".<sup>529</sup> In other words, we enjoy seeing and want to emulate the expression of just conduct to a "complex and subtle" degree. What might it mean to display just conduct in a complex and subtle manner? Rawls does not tell us directly but elsewhere he says that one pre-requisite to just conduct is seeing things from the perspective of other agents:

[W]e must recognize that these different points of view exist, that the perspectives of others are not the same as ours. But we must not only learn that things look different to them, but that they have different wants and ends, and different plans and motives; and we must learn how to gather these facts from their speech, conduct and countenance... finally, having understood another's situation, it still remains for us to regulate our own conduct in the appropriate way by reference to it.<sup>530</sup>

Rawls here refers to the ability to see things from another's perspective in order to discern how to treat them. The term "interpretive labour," which we first used in chapter one, is a good label for this capacity. It seems plausible that it could be continuously developed to call upon an ever "larger repertoire of more intricate and subtle discriminations" i.e. developed to a complex and subtle degree. If interpretive labour is necessary to just conduct and the capacity for it can be developed to a complex and subtle degree, then it seems to follow that the complexity and subtlety of just conduct is a function of the degree of interpretive labour involved in its performance.<sup>531</sup>

---

<sup>529</sup> (Rawls 1999, 413)

<sup>530</sup> (Rawls 1999, 410–11)

<sup>531</sup> There is, of course, more to being just than seeing things from another's point of view. There are other dimensions along which we could judge (and admire) the justness of a person's conduct. One



Action contexts in which discerning how one should treat others is straightforward would presumably not need a high degree of interpretive labour. One just acts according to a simple rule or convention. For example, it is surely just to give up one's train seat for someone who is clearly physically infirm. But neither does one have to think too deeply about that person's needs to realise that that is the right thing to do. Such just conduct does not appear to be to a subtle or complex degree. Higher degrees of interpretive labour would be required in those action contexts where discerning how one should treat others is *not* straightforward. Here are two examples of such contexts. We can imagine a teacher who gives appropriate dispensation to the special needs of some of his students. Perhaps, for instance, some children have a condition which means they find it more difficult to keep quiet in class compared with others. In light of this our teacher is comparatively more lenient with them if they have outbursts in class. There is a complex judgement call here because departures from uniform rules for all students arguably require justification. The teacher has to be able to discern when the particular needs of some students are pressing enough to justify such a departure. We might imagine also a senior figure in an organisation who has to decide how funding will be allocated between different groups within it. Each group within the organisation has competing demands. She determines a financial settlement all groups consider fair or at least fair *enough*. In so doing she has to consider the expectations of the different groups and the reasonableness of their demands in light of the overall aims of the organisation.

These are both situations in which individuals exhibit high degrees of interpretive labour because they have to make intricate and subtle discriminations between the needs of different parties when determining how to treat them. In such action contexts, then, it does seem appropriate to think of just conduct being expressed in a subtle and complex manner in virtue of the high degree of interpretive labour they call for. By extension, on Rawls' score, it is these forms of just conduct that are likely to elicit the admiration of others. However, this raises the following question: under what conditions are other agents suitably positioned to *recognise* such forms of moral conduct? For surely we have to first recognise instances of exemplary just conduct before we can try to *emulate* them. It seems that, at the very least, we need to know of the person in question, the circumstances in which their moral conduct

---

might be their willingness to comply with duty regardless of the personal cost. A person who is comparatively more willing to do so than another person can also be described as comparatively more just than that person. However, such a willingness to bear the burden of duty is not necessarily the same thing as exhibiting just conduct in a *complex* or *subtle* way. It is the latter dimension of evaluation that matters for one who desires social union with others.

was called for and the details of their response to it. Such knowledge seems straightforwardly to be particularistic knowledge of local time and place in the Hayekian sense.<sup>532</sup> The persons best able to recognise and appreciate the exemplary behaviour of our teacher are his colleagues and students, of our organisation manager her colleagues, and so on. In other words, only in the proximate contexts of civil society associations will we have sufficient local knowledge to recognise – and emulate – exemplary just conduct.

Indeed, as we have argued throughout this thesis, a desideratum of the rules of justice in a complex society is that they economise on the need for interpretive labour. When I meet an inevitably opaque fellow citizen in a liberal society, we should both know how to interact with one another *without* too much interpretive labour. The basic rules of justice should be comparatively straightforward to learn and applicable to a wide range of otherwise discrete cases. Their application should not require “exemplary” conduct to high degrees of interpretive nuance and complexity. Rawls’ own emphasis on the value of publicity and one of his own critiques of utilitarianism are predicated on this idea. When he criticises utilitarianism, Rawls emphasises that his conception of justice is easy to understand and apply whereas the utilitarian calculus inevitably requires complex and highly speculative calculations.<sup>533</sup> If the rules of justice regulating the well-ordered society are easy to understand and apply, then its members might appreciate it when they see their fellow citizens complying. They would also be aware that such compliance does not require a high degree of interpretive labour and by extension does not constitute the expression of just conduct to a complex or subtle degree. Hence, they would not see their fellow citizens as moral exemplars worthy of especial emulation.

Rawls might object that *political leaders* and *activists* might constitute moral exemplars who, in the context of political association, exhibit very high degrees of interpretive labour in their just conduct and whose achievements are readily apparent to their citizens. In the *Law of Peoples* he places great emphasis on the role of the “Statesman” in bringing about just outcomes under difficult and highly complex circumstances. Such figures “guide their people in turbulent and dangerous times”.<sup>534</sup> His later emphasis on the Supreme Court as the “exemplar of public reason” suggests that he believes they too might be moral exemplars in this sense.<sup>535</sup> Suppose, for the sake of argument, such exemplars do abound among the

---

<sup>532</sup> Though the content of the knowledge here differs. It is rather knowledge of a particular person’s conduct rather than, say, how a piece of machinery may be better used.

<sup>533</sup> (Rawls 1999, 277–85); (Rawls 1996, 160–12)

<sup>534</sup> (Rawls 2003, 97)

<sup>535</sup> (Rawls 1996, 231–40)

political and judicial class and that this is transparent to the citizenry at large.<sup>536</sup> The problem remains that only a small subset of the population will be in the position display this sophisticated form of just conduct. The vast majority of us will not be statesmen or sit on a Supreme Court. In a representative democracy the vast majority of us are not even engaged in political or judicial activity on a day-to-day basis. So even if these moral exemplars exist, most of us are not in a position to emulate them. In their role as *citizens*, the just conduct of most residents of the well-ordered society will be limited to legal compliance with standardised rules and will not rise to such heights of moral complexity. In summary of this section, we have argued that the well-ordered society is not the preeminent emulation social union. There is no reason to think that we cannot express our free and equal nature, in the Kantian sense valued by Rawls, by acting justly in domains of life other than the political. Participation in the well-ordered society is not necessary for such expression. Further, the well-ordered society is actually a poor candidate for an emulation social union when compared with civil society associations. For the recognition of exemplary just conduct requires local knowledge that members of the latter would have of one another but members of the well-ordered society would not.

#### 4. The Critique of the Completion Interpretation

In section two we saw that the well-ordered society is supposed to be the preeminent completion social union because it allows each to participate in a more “comprehensive good than the determinate good of individuals left to their own devices or limited to smaller associations”.<sup>537</sup> While participation in a civil society association allows me to participate in the expression of my uncultivated talents only with respect to one activity, participation in the well-ordered society allows me to participate in their expression with respect to a whole range of activities. This section will offer three points in criticism of this claim.

Beginning with the first, recall the above observation from section two that being a mere *spectator* does not allow one to participate in the expression of one’s latent talents by others. One has to play a causal role in that expression. That is what allows one to relevantly *participate* in the expression of the other person’s talent. The first criticism is that in an

---

<sup>536</sup> As opposed to adopting, say, Lord Acton’s rather more critical stance on the idea of “Great Men” in politics and ascribing the admiration of the population to false consciousness.

<sup>537</sup> (Rawls 1996, 390)

impersonal society, one's causal responsibility for the achievements of another citizen is very tenuous. Consider the following scenario:

Asif is a citizen of the well-ordered society of Great Britain. Asif was born and lives in Glasgow. When he was young, Asif showed great prospective talent in both tennis and painting. As it happens, he opted to pursue a career in painting. He remains a great fan of tennis as an adult and every year attends the Wimbledon tennis tournament in London. Sarah is also a citizen of Great Britain. She was born and raised in Cardiff. She is a professional tennis player. When he sees her compete at Wimbledon, Asif is impressed by her prowess at the game. It gives him a glimpse of what could have been. Asif and Sarah are not personal acquaintances. He had not known of her before she became a tennis player. She still does not know of him.

For sure, Sarah is expressing to a high degree Asif's latent talent for tennis which he had to forego when he opted to become a painter. Can we say, then, that Asif and Sarah are two participants in the completion social union of the well-ordered society as might be two musicians in the completion social union of a band or orchestra? The answer is negative because the disanology between this case and that of our heavy metallers Rosie and Jim seem quite clear. Rosie and Jim play causal roles in the expression of one another's talents because they play in the same band. We shall return to this point below. Asif, on the other hand, does not appear to play any causal role in Sarah's expression of her talents in tennis. He is not playing tennis with her nor is he involved in her training team in any way. He is merely a spectator in a large audience. She is unaware of his particular presence in the crowd which has no impact either way on her performance.

Rawls would probably respond that Asif in his capacity as a citizen has played a causal role in Sarah developing her tennis skills. I here draw upon Paul Weithman's plausible exegesis of why Rawls might believe this to be the case.

If members of the [well-ordered society] do their part in upholding just institutions, and they know that just institutions make it possible for others to pursue their good—whether it be baseball or music, stamp collecting, or family life—then each person can see the pursuits of others as developments of her own latent abilities that she has helped to make possible.<sup>538</sup>

There are two sequences in this causal chain. Applied to our above scenario, we might say that (i) Asif helps maintain just institutions by doing his part as a dutiful citizen; and (ii) those just institutions enable Sarah to become a world-class tennis player. Both causal sequences are tenuous at best. Let us overlook the analysis of chapter two and assume for the sake of argument that Great Britain has a basic structure which accords with justice as fairness or

---

<sup>538</sup> (Weithman 2011, 139)

some other reasonable conception of justice. Asif is but one of millions of British citizens whose actions have contributed to maintaining just institutions. His own causal contribution to the justness of British institutions would be negligible. Suppose it turns out that Asif is actually a serial tax evader, his marginal defection would have no causal impact on the justness of British institutions.<sup>539</sup>

The causal connection between Britain's just institutions and Sarah's successful tennis career is also questionable. Supposing that Sarah has a natural talent for tennis, it seems likely that she could also have become a tennis player under a variety of alternative regimes. Rawlsian institutions do not seem *necessary* for Sarah to manifest her tennis skills. Perhaps the claim is that the well-ordered society has enabled Sarah to develop her tennis capacities to *greater degrees* of complexity and subtlety than she would have done otherwise. However, it is not hard to imagine alternative institutional regimes under which she would have developed her tennis skills to an even greater degree than she did under the Rawlsian regime. Economic inequalities are allowed in the well-ordered society only to the extent permitted by the difference principle. The income one earns from playing top level tennis would likely be much less than under a more free market regime. Suppose that an agent's propensity to invest time and effort in an activity depends at least partly on the monetary awards. Then perhaps Sarah would have become an even more proficient tennis player under a free market regime in which she could have kept more of her income.<sup>540</sup>

It is even plausible that she would have become a more proficient tennis player under an authoritarian regime. Suppose that instead of being a well-ordered society, this hypothetical Great Britain was an authoritarian regime that promoted sport for propaganda reasons.<sup>541</sup> The regime might pick tennis as one of the key sports in which Great Britain would attain national greatness. Sarah might have become an even better tennis player under such a regime because she would have received greater funding and institutional support compared with the well-ordered society. Further, the scenario above was predicated on the idea that Sarah and Asif are both citizens of Great Britain. However, the displays of human excellence

---

<sup>539</sup> The Rawlsian might say he is acting *unfairly* in not doing his part but that is not the same as *jeopardising* the outcome.

<sup>540</sup> The Rawlsian would object that this means she is motivated by money rather than developing her capacities for their own sake as the Aristotelian principle implies. Not necessarily. Rawls himself assumes that monetary incentives are needed to cover things like training costs, and becoming a tennis player (I conjecture) is not cheap. Alternatively, even a Rawlsian might not object to someone picking the more remunerative option to develop out of two or more capacities she finds equally intrinsically rewarding.

<sup>541</sup> As, say, the Soviet Union used to do.

to which we are exposed and in which we take pleasure in a liberal society are hardly going to be limited to only those of our co-nationals. Indeed, Wimbledon is renowned above all for its internationalism.<sup>542</sup> We can imagine an amended version of the above scenario in which Sarah was a player representing another nation. Under such a scenario there seems no link whatsoever between Sarah's development of her tennis skills and British institutions or indeed Asif's conduct.<sup>543</sup>

While the causal connections between Asif's compliance with justice and Sarah's expression of her tennis skills is quite tenuous, matters are different with our heavy metallers, Rosie and Jim. When Rosie and Jim play "Master of Puppets," Jim can be confident he plays a causal role in her performing her part. For Rosie is playing the bass guitar to "Master of puppets" in part *because* Jim is playing drums. There is usually no point to playing a rock song without the percussion. Now, if we restrict ourselves to the band playing a song, it seems that we have shown Jim to be causally responsible only for the expression of Rosie's talents on this *one occasion* but not its *development* over time. However, we could easily amend the scenario and imagine Rosie and Jim to be band mates who have rehearsed with one another for years. In which case, we might suppose that frequent rehearsals with Jim helped Rosie develop her bass guitar skills. Hence, Jim would have played a tangible causal role also in their development. (And vice versa.) The point is that the causal connections here are comparatively clear and straightforward.

Our analysis parallels the Hayekian claim that we saw in chapter two about the complexity of the order of actions that we find in liberal societies. In small-scale social orders, where the participants are few in number, we can understand the interrelations between the actions of the participants. This is certainly true of our heavy metal band. If one of the players is out of time, for instance, the others will notice. Matters change under large-scale cooperative schemes consisting of a mass of actors operating under conditions of fragmented knowledge. As we just saw, the difficulty is not simply that the sheer number of actors makes it difficult to allocate a causal role to any one of them in particular, especially in instances where the aggregate outcome is over-determined. It is also that the causal connection between individual actions and the resulting order of actions become obscure and highly conjectural

---

<sup>542</sup> There is an even a term, the "Wimbledon Effect," for the British propensity to host competitive institutions in which the prominent players come from abroad. The term has been used to describe the City of London.

<sup>543</sup> I appreciate that Rawls models the well-ordered society as a closed society. Since the only actual society that seems to meet this ideal is North Korea, I submit it to be a dubious modelling assumption.

at best. It seems clear if not trivially true that Rosie contributes the bass line to her and Jim's performance of "Master of Puppets" in part because Jim is playing the drums. It seems far more dubious to say that Sarah is playing tennis at Wimbledon – even in part – because Asif obeys the law and pays his taxes. If the causal contribution of liberal citizens to the expression of one another's talents is tenuous, the well-ordered society might not be a completion social union at all, let alone the preeminent one. For liberal citizens do not relevantly *participate* in the expression of one another's talents. They are more akin to spectators.

Let us move onto the second point. This can be stated rather more briefly. Most of us have but a handful of latent talents. Few of us are polymaths. If so, and setting aside the question of whether we really participate in the expression of others' talents in the well-ordered society, most of the diversity of talents on display in it will not help to complete our "selves" in the way Rawls and von Humboldt value. They will simply be wasted on us. Perhaps Rawls can reformulate his conception of how the self is completed in response to this point. The experience of self-completion might not come only from seeing others express talents that I latently *have* but chose not to develop. It might come also from seeing others express talents that I simply *do not have* but *wish* I did.<sup>544</sup> This would allow us to feel "completed" by more of the diverse talents on display in the well-ordered society than we might do otherwise. Especially those of us whose range of latent talents might be somewhat meagre. This rendering of the claim is more plausible but it remains the case, I think, that there are only so many capacities that one *wishes* they had. I certainly wish I had a latent talent for playing the guitar. In fact, I most certainly do not. Perhaps I can still feel "completed" in some sense when I hear Jimi Hendrix express my yearned-for but altogether lacking capacity for guitar playing. Still, I do not feel this way about a host of activities that one finds in an open liberal society. Just to take some ad hoc examples, I am confident I also have no latent talent in either playing darts, performing ballet or competitive sailing. However, nor do I have any particular wish that I had such latent talents. The expression of these and many other competences on display in an open and liberal society would presumably, then, not aid in the completion of my "self". Another way of putting this point is to say that once we recognise that there are only so many latent or even wished-for talents one could have, the "value-added" of participation in the well-ordered society in completing the self seems quite

---

<sup>544</sup> Many of us, I think, feel inadequate or "incomplete" because we lack capacities that we wish we had rather than because we opted not to develop capacities we knew we had.

limited. If we each only really have a handful of talents, it seems that one can adequately complete the self by participating in a handful of civil society associations.

Moving onto the third point, even where the talents on display do represent the cultivation of our latent abilities, it is not clear we would have the knowledge to recognise and appreciate the subtlety of such displays. Recall that per the Aristotelian principle, it is the display of capacities to a high degree that are enjoyed by the self and others. The problem, however, is that one often has to be a seasoned practitioner in an activity themselves to appreciate its expression to a high degree of complexity. This seems true of Rawls' orchestra or our heavy metal band. Rosie might not be a drummer herself but she can recognise Jim's excellent drumming because she has general musical expertise. Suppose now she has a latent capacity in art but one that is relatively undeveloped. While she may have done well at art class in school she has not cultivated the skill since and spends little time with artists. She (and other liberal citizens) face a peculiar kind of knowledge problem in recognising excellences in fields where they have only undeveloped talents.

To make this particular point, I wish to draw on an insight of Alasdair MacIntyre. There are some striking similarities between Rawls' account of a social union and MacIntyre's account of a social practice as expounded in his *After Virtue*. This should not be too much of a surprise since both draw on Aristotle. MacIntyre defines a practice as

... [A]ny coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and goods involved, are systematically extended.<sup>545</sup>

Like members of Rawls' social union, MacIntyre's practitioners should be intrinsically motivated. For example, to be properly engaged in the practice of chess, I should want the "goods internal to the practice of chess which cannot be had in any way but by playing chess or some other game of that specific kind".<sup>546</sup> I am not properly dedicated to chess if I am in it for the money or adulation since these extrinsic goods can be obtained by participation in any kind of activity.

The similarity between the two accounts established, it is worth turning to one part of MacIntyre's account of a practice that is surely also true of a social union, but about which

---

<sup>545</sup> (MacIntyre 2007, 187)

<sup>546</sup> (MacIntyre 2007, 188)



Rawls places comparatively less emphasis. MacIntyre observes that inculcation into a practice presupposes that one recognise their ignorance of the prevailing standards of excellence in that practice:

To enter into a practice is to accept the authority of those standards and the inadequacy of my own performance as judged by them. It is to subject my own attitudes, choices, preferences and tastes to the standards which currently and partially define the practice.<sup>547</sup>

MacIntyre's point is not simply that the novice has to acknowledge that their performances fall short of the prevailing standards of excellence in the practice. The novice has to also acknowledge their ignorance of what those standards indeed are and their inability to *recognise* the manner in which already competent performers meet those standards.<sup>548</sup> A large part of developing one's latent capacity for an activity is becoming acquainted with the standards of excellence in that activity about which one knows little beforehand. Thus, one cannot recognise excellence in portrait painting unless one has substantial experience of painting portraits. Rawls, I think, would concur with MacIntyre's claim here. Recall that his Aristotelian principle holds that, all other things being equal, we enjoy performing those activities that call on a "larger repertoire of more intricate and subtle discriminations". It seems plausible to say that there is no hard-and-fast distinction between being able to recognise such intricate and subtle discriminations in the work of others and being able to exhibit them in one's work. They are part-and-parcel of the same process.<sup>549</sup>

It follows that one's ability to recognise the expression of excellence by others in a field in which they have only latent or undeveloped talent will be limited. They will not have the detailed and practice-specific knowledge to properly recognise the excellence-in-question. A latent talent for mathematics will not allow one to understand the latest advances in the discipline. Indeed, the nature of specialisation is such that even within the same discipline, practitioners of different *sub*-disciplines might not be able to understand the nature of one

---

<sup>547</sup> (MacIntyre 2007, 190)

<sup>548</sup> Thus he goes on to write: "[if], on starting to listen to music, I do not accept my own incapacity to judge correctly, I will never learn to hear, let alone to appreciate, Bartok's last quartets" (MacIntyre 2007, 190). The point here, I take it, is that the novice cannot *recognise* the standards of excellence displayed in a competent performance of a Bartok quartet and not simply that they would *fail to meet* those standards were they to play a part in a Bartok quartet themselves.

<sup>549</sup> I should say I think MacIntyre's claim is put too strongly. Surely the *connoisseur* can recognise great art or a fine wine without being an artist or a winemaker. Still, even connoisseurship requires the investment of time and effort to gain the requisite knowledge of the standards-in-question. But we can still affirm MacIntyre's basic claim, which surely has common sense on its side, that the ability to recognise excellence in a field requires the cultivation of specialist ability, whether that ability is that of the direct practitioner or the connoisseur.

another's contributions. This is certainly true of say philosophy. We need only think of the mutual miscomprehension that often exists between so-called analytical and continental approaches to the topic. This knowledge problem here is similar to that of Hayek but also importantly different. His problematic pertained to knowledge of particular facts of time and place. In an economic context, it is the knowledge of the use that might be made of a particular resource or how, say, a piece of machinery may be used in a specific way. This knowledge was spatially and temporally distributed in such a way that it cannot be centrally aggregated. Now, knowledge of the standards of excellence pertaining to different human disciplines or practices are not spatially and temporally distributed in the above sense. Presumably, the art expert can recognise great art wherever and whenever she sees it. Nonetheless, what is similar is that such knowledge is dispersed and cannot be aggregated. It is dispersed because, due to scarcity of time and also to personal disposition, each can only come to know what true excellence looks like in at most only a handful of fields.

In summary of this section, we have criticised the claim that the well-ordered society is the preeminent completion social union. Rawls' claim being that participation in the well-ordered society allows one to realise a much greater array of their latent talents than they would by participating in smaller civil society associations. We have offered three criticisms of this view. First, it is not clear that the well-ordered society qualifies as a completion social union at all since citizens do not relevantly participate in the expression of one another's talents. Second, there are only so many latent talents or otherwise wished-for talents that one might have. The added value, with regard to completing the self, of the sheer variety of talents on display in the well-ordered society is questionable. Third, and in any case, our ability to *recognise* excellent expression of our latent talents in practices in which we lack practical experience is highly limited.

## **Conclusion**

This chapter has criticised the Rawlsian claim that the well-ordered society can be understood as a social union. Here we can take a step back and undertake some more general reflections, comparing Rawls' understanding of a liberal society as a social union with the very contrasting approach offered to us by Hayek. The completion model of the social union is here assumed to be the authoritative one. In his invocation of the well-ordered society as a social union, Rawls asks us to imagine a society where liberal citizens are mutually

aware of how their discrete exercises of freedom cohere into a wider harmony. This is, after all, how it feels in his ideal form of a social union, the orchestra. Each player can perform their part with at least some degree of individual flair and each has consciously chosen which instrument to play. Nonetheless, their individual performances cohere into a wider whole all can survey and make sense of. Indeed in §79 he writes that in the social union of the well-ordered society, “[everyone’s] more private life is to speak a plan within a plan, this subordinate plan being realized in the public institutions of society”.<sup>550</sup> The nub of our critique in this chapter has been in line with the general thrust of this thesis. We have accepted for the sake of argument that all citizens of the well-ordered society subscribe to the conception of goodness as rationality and strive for social union. We have focused on identifying epistemic reasons as to why the well-ordered society as a whole should not be regarded as a social union. The emergent order of actions and its workings are too opaque for its many participants to see themselves as participating in the expression of one another’s talents. As in the previous chapters, we have emphasised that it is mutual ignorance rather than disagreement that prevents the mass of liberal citizens from acting-in-unison.

Hayek, like Rawls, stresses that we benefit from the talents of others in a liberal society. However, his emphasis on human ignorance leads him to construe this claim rather differently. In the Rawlsian social union we are *consciously aware* of how we are enriched by the talents of others. Hayek, by contrast, emphasises how in the Great Society we are enriched by the capacities of countless others in ways that we can scarcely understand. We can recognise the start of “civilisation” when “the individual in the pursuit of his ends can make use of more knowledge than he himself has acquired”.<sup>551</sup> As we observed in chapter one, on the emergent view one of the functions of liberal institutions is to transmit otherwise unavoidably disaggregated forms of knowledge. The price mechanism transmits dispersed economic information. More generally, decentralised processes of social evolution and competition allow for the diffusion of successful cultural practices via mechanisms of imitation. Very often those who adopt the more successful practices will not know *why* they work. They simply know *that* they work. The point here is that, for Hayek, the liberal order allows us to take advantage of the capacities of countless other persons. Such capacities *far exceed*, rather than purport to complete, the latent capacities of the singular self.

---

<sup>550</sup> (Rawls 1999, 463)

<sup>551</sup> (Hayek 2006 [1960], 21)

Relatedly, one of the striking aspects of Rawls' thought generally but one that manifests itself particularly in his thinking of social unions is the lack of emphasis on *social learning*. The individual knows all her capacities but alas can secure only a subset of them. Participation in the liberal community allows her to realise these latent capacities via the efforts of others thereby avoiding the perils of partial cultivation. It is hard to make sense of this notion without at least implicitly assuming there is a fixed set of latent capacities or ends that each person has to secure in life. Suppose, instead, that the self could develop new capacities. Then it is not clear what it means to "complete" the self. For even if I manifest all the capacities that I find to be of value at a particular moment in time, new ones could subsequently arise and my "self" would no longer be "complete". Such is Hayek's view. Our understanding of value and what our capacities consist in are not "givens" but subject to processes of social learning. When others utilise their particular knowledge to discover new and creative practices, exposure to them could *change* our understanding of both what our capacities might be and why they are to be valued.<sup>552</sup> On Hayek's score, individual freedom is to be valued insofar as it facilitates such experimentation and learning. He even claims that if we could predict the results of widespread exercises of freedom, it would have no value.<sup>553</sup> The emphasis of his liberalism, then, is not on the completion of the self but its ability to learn new things. The following citation, then, seems a fitting way to end this passage and chapter:

It is not in the fruits of past success but the living in and for the future in which human intelligence proves itself. Progress is movement for movement's sake, for it is in the process of learning, and in the effects having learned something new, that man enjoys the gift of intelligence.<sup>554</sup>

---

<sup>552</sup> Such is the theme of chapter 2 of the *Constitution of Liberty* (2006), on which this discussion has drawn.

<sup>553</sup> (Hayek 2006 [1960], 27–28)

<sup>554</sup> (Hayek 2006 [1960], 38)

## Chapter Seven: Alleviating Hayek's Worry

### Abstract

This thesis has argued that collective political agency is not possible because of the mutual ignorance obtaining between liberal citizens. However, Hayek worried that the desire for collective control of the social order is hard to dislodge because it is an evolutionarily inherited intuition. While such control can never be attained, the belief that it is attainable will lead to forms of political activity that undermine liberal institutions. This chapter alleviates Hayek's worry by making two claims. First, even if we grant an innate human preference for a purposive and fraternal social order, our tendency to express that preference in politics is not an invariant social fact. Such a tendency is at least partly endogenous to the political institutions under which individuals are socialised. Second, liberal institutions in which authority is dispersed and whose operation inhibit the formation of faction may be prove robust even if many of their citizens are attracted to collectivist or otherwise illiberal political projects.

### Introduction

The first chapter of this thesis distinguished between the emergent and associational views of political society as represented in the liberalisms of Hayek and Rawls respectively. These views are premised on different accounts of the manner in which liberal citizens are alienated from one another and prevented from acting-in-unison in the pursuit of shared aims. The associational view assumes liberal citizens are alienated by mutual disagreement and the emergent view by mutual ignorance. The first chapter argued that the prospects for collective political agency seem quite sanguine if one assumes mutual disagreement to be the cause of alienation. One then just needs to identify impartial principles over which liberal citizens can agree. Its prospects seem much less sanguine, however, if mutual ignorance is the cause of alienation. This is a much deeper constraint on collective political agency. On this view, the mass of liberal citizens simply do not know enough about one another to act-in-unison. Large-scale political institutions necessarily are the emergent result of many disparate human actions and not the expression of intentional collective agency. The previous chapter stressed that, for the same reason, the individual cannot contextualise her own achievements as a contribution to a greater whole.

If the argument thus far is successful, it also brings into full view Hayek's worry about the stability of liberal institutions which we first saw in chapter one. Hayek fears that the desire for collective political agency is the expression of an atavistic evolutionary intuition that we

cannot easily discard. We are constantly prone to apply the “unmodified, uncurbed, rules of the micro-cosmos (i.e., of the small band or troop, or of, say, our families) to the macro-cosmos (our wider civilisation)”.<sup>555</sup> Political projects promising to reconcile the “macro-cosmos” with the unity and collective control of the small tribe will prove irresistible. While such collective political agency will never be achieved for epistemic reasons already outlined in this thesis, the *attempt* to achieve it will encourage modes of political activity ultimately incompatible with a free society. Or such is the thesis that Hayek offers principally in his final work *The Fatal Conceit* but also in parts of *Law, Legislation and Liberty*. It now behoves this thesis to give full consideration to this worry. This is the task of this final chapter. It will try to alleviate Hayek’s worry by showing that our demand for a purposive and fraternal order is not as invariant as the worry assumes and that in any case a liberal order can prove robust even if not subject to the deep endorsement of the governed. We can happily draw on some of Hayek’s other ideas, as expounded in other parts of his corpus, in this task of alleviation.

This chapter is divided into five sections. The first outlines Hayek’s worry in greater detail. It shows why Hayek thinks the demand for collective agency undermines freedom in liberal democracies. The remainder of the chapter tries to alleviate Hayek’s worry. This can be done by way of two argumentative strategies. One might be to show that the belief that the liberal order should be subject to collective control is more variable than Hayek’s worry assumes. Another would try to show that there are conceivable liberal regimes that could endure even if large swathes of their population would prefer a collectivist regime. This chapter pursues both strategies. Thus, the second section argues that the belief that the liberal order should be collectively controlled is not simply contingent on our evolutionary inheritance but also socially conditioned. There might be conceivable social worlds where citizens do not hold that belief as strongly as they do under existing social democratic regimes. Sections three and four articulate a conception of liberal stability in which institutional robustness is an emergent phenomena arising from the interactions of very many agents who individually may not intend that result. This conception of liberal stability is presented as one that is conducive to the emergent understanding of political society. I term it *emergent stability*. This understanding has a lineage going back to David Hume and James Madison. A liberal regime utilising emergent stabilising forces could endure even if large swathes of its citizens were attracted to collectivist political projects since the resulting stability would be the

---

<sup>555</sup> (Hayek 1988, 14)

unintended outcome of their various interactions. Section five responds to a key Rawlsian objection to the use of emergent stabilising forces.

### 1. Outlining Hayek's Worry

As already outlined in chapter one, Hayek conceives of the liberal order as a "Great Society". It consists of many different individuals, groups and communities who act upon their own contextual knowledge and understandings of value. The unity of such an order does not extend beyond subscription to common rules of just conduct.<sup>556</sup> Both the institutional rules underlying this complex system and the order of actions to which they give rise will not be amenable to conscious collective control and regulation. Our moral intuitions, however, evolved in much smaller social orders characterised by very much higher levels of mutual knowledge between a very much smaller number of participants. We are evolutionarily prone to make the category error of expecting social relations in the Great Society to take place on the same purposive and fraternal terms as that of the small troupe. This tendency threatens to undermine a free society.<sup>557</sup> When Hayek first engages in political theory, the popular collectivist project was socialist central planning. Hence in *The Road to Serfdom* he argues that socialism is incompatible with freedom. Later in his career, when the popularity of socialism fades, Hayek thinks collectivist politics expresses itself as the demand for democratically enforced social justice.<sup>558</sup> Hence in *Law, Legislation and Liberty*, Hayek criticises social justice as incompatible with freedom. I take it to be clear how socialist central planning undermines liberal freedom. That is uncontroversial even among left-liberals who are otherwise sceptical of the market economy. It is worth explaining why Hayek thinks the demand for social justice also undermines freedom. He worries that the widespread *belief* in social justice provides cover for interest group political activity. The cumulative effect of such political activity is the building up of rules and regulations that undermine freedom.

In any moderately complex society individuals will form interest groups and engage in politics to secure benefits for themselves or otherwise secure goals they value using the machinery

---

<sup>556</sup> In chapter two we suggested that even this might be an over statement insofar as a complex society will have multiple makers and enforcers of institutional rules.

<sup>557</sup> See particularly Hayek (1988, chap. 1) and Hayek (1982b, chap. 11).

<sup>558</sup> Which, as we saw in chapter two, Hayek understands primarily as a patterned conception of distributive justice.

of the state. There is a natural, and for Hayek healthy, suspicion that some groups use state machinery to further their own partisan ends and not goal that the wider populace might endorse. A widespread belief in social justice undermines this suspicion. It encourages the thought that the claims of various interest groups can be reconciled with a discernible standard of the common good shared by all citizens. Of course, Hayek thinks this mistaken. In an impersonal order of agents mutually ignorant of one another's circumstances and values, no such standard for the common good exists. This is one of Hayek's key claims in the second volume of *Law, Legislation and Liberty*. The problem is that the belief in social justice provides ideological cover for such interest group politics. The expectation that, to coin a phrase, society is a cooperative venture between free and equal citizens who work together to realise social justice legitimates and encourages the rent seeking of special interests. For interest groups or their elected representatives can then draw on the widely accepted discourse of social justice to present – both to others and perhaps even to themselves – their partisan and private demands as being in the public interest. As Hayek puts it in the third volume of *Law, Legislation and Liberty*:

The creation of the myth of 'social justice' which we have examined in the last volume [i.e. volume 2, *The Mirage of Social Justice*] is indeed largely the product of this particular democratic machinery [i.e. interest group politics], which makes it necessary for the representatives to invent a moral justification for the benefits they grant to particular interests.<sup>559</sup>

Still, it might be asked, the propensity to form factional interest groups and engage in rent seeking might be economically inefficient but it is not clear how it might threaten *freedom*. Hayek, I don't think, doesn't fully explain his thinking here so we need to explicate them a bit. Under simple majority voting rules, 51% of the population are able to impose their preferences on the remainder. The problem here is that electorally successful political coalitions engage in political transactions the costs of which are shared by everyone in society.<sup>560</sup> Suppose the deal between the current governing coalitions involves, say, restrictions on smoking and extra subsidies for farmers. The losers in the election would both have to contribute to the subsidies and comply with the smoking restrictions (or face the

---

<sup>559</sup> (Hayek 1982c, 10) c.f. "Yet today, if ... [an interest] group is only large enough, its representation of the demands of its members as just is commonly accepted as one view of justice which must be taken into account in ordering the whole, even though it does not rest on any principle which could be generally applied" (Hayek 1982b, 39).

<sup>560</sup> I draw here from chapter 3 of *The Calculus of Consent* of Buchanan & Tullock (1999), whose analysis has much in common with that of Hayek in the third volume of *Law, Legislation and Liberty*.



costs of non-compliance) despite not having supported the policies.<sup>561</sup> Over time, as various political coalitions gain the chance to govern and fulfil the wishes of their constituents, restrictions on the liberties of all will build up cumulatively. Hence the deeply managerial nature of contemporary liberal democratic states. There are few areas of modern life that are not subject to some kind of regulation and control by state actors and these inevitably reduce the scope for individual decision-making.<sup>562</sup> In sum, the expectation that society is a collective venture realising social justice legitimates interest group politics. The unintended consequence of such politics is the accumulation of laws that undermine liberty across various dimensions. Our reconstruction of Hayek's worry completed, we will now move onto showing how it might be alleviated.

## 2. The Endogeneity of Political Expectations

Our evolutionarily inherited belief that the liberal order should be subject to collective control might be tempered by other factors. Hayek's other work can help us identify three:

- a) The prevailing worldview popularised by the intellectual class
- b) The legitimisation strategies employed by political institutions
- c) The vibrancy of civil society

We shall look at each of these in turn.

In his essay "The Intellectuals and Socialism", Hayek argues that the popularity of socialist ideas among the general populace is a function of the ideas prevailing among the intellectual classes.<sup>563</sup> "Socialism" is here understood as the centralised planning and arrangement of society to secure common goals.<sup>564</sup> For Hayek the intellectual class consists not only or even primarily of academics but also of members of any profession who act as intermediaries between the intellectual world of ideas and the public-at-large, ranging from journalists to

---

<sup>561</sup> By contrast, regular market exchange does not allow one to coerce others in this manner.

<sup>562</sup> Note that, consistent with our analysis of state institutions in chapter two, these restrictions are themselves the emergent outcome of many interactions between different interest groups. There is no central agency primarily responsible for them.

<sup>563</sup> (Hayek 1949)

<sup>564</sup> Hence, he by no means saw socialism as restricted to the left and dedicated *The Road to Serfdom* to the "socialists of all parties".

radio commentators to artists.<sup>565</sup> He offers what we might term a “trickle-down” model of the battle of ideas. While intellectuals have little short- or medium-term impact on public opinion or political decision-making, their long-run influence is considerable. Intellectual views that are ascendant among the intellectual class in one generation will be popular opinion the next and influence politics. One changes politics by first winning over the intellectuals. Hayek’s goal in the essay is to persuade his fellow classical liberals to concentrate on winning the intellectual battle of ideas so as to change public opinion at large in the long-run. As Kukathas observes, Hayek himself dedicated much time and energy to this endeavour, from advising Sir Anthony Fisher to establish The Institute of Economic Affairs to himself helping establish the Mont Pèlerin Society.<sup>566</sup>

For the purposes of this chapter, we can infer the following. Even if our inherited moral intuitions instil a deep and perhaps emotive preference for the concrete and face-to-face relations of the small group, our ability to discipline this emotive preference when formulating our expectations of large-scale political institutions would also be influenced by the prevailing world view of the intellectual class.<sup>567</sup> If the intellectual consensus informs us that it is feasible to expect of large-scale political institutions the same expectations apropos of small-groups, then of course we are unlikely to engage in such discipline. However, this is arguably a contingent fact. Indeed, in his paper Hayek observes that socialist ideas gained ascendancy among the intellectual class only in the 20<sup>th</sup> century in Britain and the US and only in the late 19<sup>th</sup> century in Germany.<sup>568</sup> The attraction of the public to such political ideas increased subsequent to these shifts. Similarly, in *Law, Legislation and Liberty* Hayek’s thinks the respectability conferred to social justice by *intellectuals* makes it much more attractive to the population at large than it otherwise might be.<sup>569</sup> The converse might be true by the same token. If the prevailing intellectual worldview shifts towards a classical liberal direction, we should expect a similar shift in public opinion.

We now move to the legitimation strategies employed by political institutions. This endogeneity of political beliefs to institutions was one of Hayek’s key concerns in *The Road to Serfdom*. That work is often interpreted as advancing what Meadowcroft and Alves have

---

<sup>565</sup> (Hayek 1949, 418)

<sup>566</sup> (Kukathas 2006)

<sup>567</sup> (Hayek 1949, 423)

<sup>568</sup> (Hayek 1949, 417)

<sup>569</sup> (Hayek 1982c, 135–36)

termed the slippery slope thesis that the interventionism of the mixed economy will lead to totalitarianism.<sup>570</sup> The mechanism is that the failures of state intervention beget yet more state intervention until the mixed economy turns into a centrally planned one. Such a sociological thesis now seems implausible. Whatever their other faults, social democracies have seemingly resisted this inexorable shift towards Soviet-style planning, leading even classical liberal sympathisers of Hayek in Meadowcroft and Alves to declare the argument a failure.<sup>571</sup> No doubt, something like the slippery slope thesis can be read into *The Road to Serfdom*. However, in the preface to the second edition of the book Hayek laments that his most important argument had been overlooked, that the “political ideals of a people and its attitude towards authority are as much the effect as the cause of the political institutions under which it lives”.<sup>572</sup> What concerned Hayek about socialism was the kind of subject it might produce – “the most important change which extensive government control produces is a psychological change, an alteration in the character of the people”.<sup>573</sup>

Comprehensive economic planning would require that many otherwise individual decisions be taken collectively. A regime engaged in such planning would try to encourage and cultivate our evolutionary instincts for a legible and purposive social order for reasons of its own legitimacy. It would encourage us to think of society as a very large organisation in which one’s station in life is the function of a collective, political decision. If it succeeded in doing so the citizenry would be likely to accept planning in part because they would be socialised to do so. Now, *The Road to Serfdom* is concerned with the modern state in its most extreme and worrisome manifestation of comprehensive socialist planning. Still, the psychological thesis arguably has force for even contemporary social democratic states. While their propensity to regulate society is much reduced when compared to centrally planned regimes, it is nonetheless much greater than, say, the typical European state in the nineteenth century. It seems plausible that citizens socialised in such social democratic regimes would be disposed to see one another as participants in a democratic project to realise shared standards of social justice. Conversely, we might expect individuals socialised under a more classical liberal regime to be less disposed to see political life in this way.

---

<sup>570</sup> (Alves and Meadowcroft 2014)

<sup>571</sup> See particularly Alves and Meadowcroft (2014, 854-856). For an argument that the slippery slope thesis is more nuanced and defensible than it might initially appear, see Cowen and Tranditis (2019).

<sup>572</sup> (Hayek 1967c, 224)

<sup>573</sup> (Hayek 1967c, 224)

There are some clear parallels here between Hayek's observations and those of Michel Foucault. The latter observed that the "art of government" under modernity differs drastically from the past. In the Middle Ages, sovereignty is exercised not on "things" but rather on a territory.<sup>574</sup> By this Foucault means that the aim of governance was primarily the control of territory and not so much the regulation of the individuals and communities who resided within them. Such an art of governance is presumed within Machiavelli's *The Prince*, "essentially a treatise about the prince's ability to keep his principality".<sup>575</sup> Indeed, while in *The Prince* Machiavelli does offer what one might think of as managerial advice, this is with respect to the Prince's personal estate. Machiavelli offers advice on what kind of staff to appoint or the wisdom of building castles for defence or of disarming the population. But these forms of managerial advice are instrumental to securing the Prince's personal goals of maximising the territory under his control. The *population itself* is not the subject of management.<sup>576</sup> The manner in which the art of governance is conceptualised begins to change, however, as we approach the modern era. Now, one governs *things*. The focus moves away from the territory itself to the communities within them and the regulation and governance of their social relations.<sup>577</sup> This contemporary art of governance is what Foucault would term *governmentality*. What makes such governance possible is the development of statistics. These reveal the population of a territory to show aggregate level regularities across a variety of dimensions from wealth to health. These regularities can be identified, measured and intervened upon.

Foucault's thesis is in line with the historical sociology of the state. Van Creveld observes that the size and scope of even the absolutist states of the 18<sup>th</sup> century were very much less than their contemporary successors. Taxation was much more limited. The most heavily taxed eighteenth century state, Frederick II's Prussia, appropriated the same proportion of national income as did the United States in 1989.<sup>578</sup> The tasks of day-to-day governance were largely left to intermediary groups.<sup>579</sup> As Michael Oakeshott observes, states in the middle ages were composed of diverse and largely self-governing communities united by little more than a common ruler whose main task was to keep the peace.<sup>580</sup> Matters change in the 19<sup>th</sup> century

---

<sup>574</sup> (Foucault 1991, 93)

<sup>575</sup> (Foucault 1991, 91)

<sup>576</sup> (Machiavelli 2004)

<sup>577</sup> (Foucault 1991, 91)

<sup>578</sup> (Van Creveld 1999, 190)

<sup>579</sup> (Van Creveld 1999, 190–91) C.f. (Nisbet 1953)

<sup>580</sup> (Oakeshott 2003, 242)

when even a purportedly laissez faire state like that of Britain begins to make comprehensive efforts to intervene in society. It begins to invest heavily in crime prevention and social control. The number of prosecutions grow seven fold between 1805 and 1840 and new legislation makes it possible to prosecute individuals for mere vagrancy.<sup>581</sup> The state begins also to regulate economic life, the first Factory Act being passed in 1834. It moves into welfare provision and starts to regulate public health. Welfare is appropriated from the parishes and provided directly by the state through the infamous Victorian workhouses. The first Public Health bill is passed in 1848 giving local boards of health authority to control water supplies and maintain streets. The General Medical Council is formed in 1858 to oversee and regulate the medical, surgery and midwifery professions.<sup>582</sup> Additionally, Britain along with other 19<sup>th</sup> century states, used transport policy, particularly with respect to rail, to connect their provinces and bring them under central control.<sup>583</sup> The scale of activity performed by twentieth century states is of a greater magnitude and scope still.

Foucault's insight is that the shift in the art of government from the management of territory to the management of things necessitates a shift in the self-understandings of the governed—one that would both come to value having the goods promoted by these interventions and see the successful promotion of these goods as an appropriate mode of government activity. This is what Foucault has in mind when he speaks of the normalisation of the subject under governmentality.<sup>584</sup> Today, most of us take for granted that the function of nation-states is to promote our welfare in a plethora of ways, ranging from the education of our children to the maintenance of our health to equipping us with skills to survive in the global economy to protecting our culture and so on. When we enter the voting booth, we judge our governments on how effective they are in meeting these goals. We have been socialised into thinking of our societies as organisations governed by the state. This is a new self-understanding among subjects, one not found in the past when the size and scope of state activity was much less than it is currently. Foucault's analysis complements Hayek's view, as expounded in *The Road to Serfdom*, that our expectations of political institutions are formed not simply by our genetic inheritance but are endogenous to those institutions themselves.

---

<sup>581</sup> (Van Creveld 1999, 210)

<sup>582</sup> (Van Creveld 1999, 219)

<sup>583</sup> (Van Creveld 1999, 260)

<sup>584</sup> I follow here Diana Taylor's (2009) helpful interpretation.

We now move unto the third variable, the vibrancy of civil society. In a passage advocating radical decentralization of government services in the third volume of *Law, Legislation and Liberty*, Hayek writes that it:

... would probably lead to the revival of a communal spirit which has been largely suffocated by centralization. The widely felt inhumanity of the modern society is not so much the result of the impersonal character of the economic process, in which modern man of necessity works largely for aims of which he is ignorant, but of the fact that political centralization has largely deprived him of the chance to have a say in shaping the environment which he knows.<sup>585</sup>

In chapter four we saw Tocqueville's thesis that centralisation of social life by the state undermines the propensity for local association. The decline of worker-led friendly societies in the UK and US following greater state provision of welfare was offered as suggestive evidence in favour of the thesis. It is relevant also for the concerns of this chapter. In a vibrant civil society, our need for fraternal and purposeful social organisation could be partly met in associational life.<sup>586</sup> For instance, demands for distribution according to moral desert, inappropriate when made on the impersonal liberal order, could be made appropriately in such settings. The aforementioned friendly societies could take into account moral desert in their distribution of entitlements because the members of a particular chapter would be personally acquainted with one another. Information as to one another's conduct could be ascertained relatively easily. However, once the central state displaces these local associations, they can no longer partially meet our need for collective control and fraternity. The demand on large-scale political institutions to meet such needs will increase. By the same token, if the demand for collectivist politics is high in centralised social democratic regimes, it might be less so in regimes which afforded local associations greater autonomy.

### 3. Two Accounts of Liberal Stability

We now change tack and try to show that there are conceivable liberal institutions that might endure even if large swathes of their subjects would prefer to overturn them in favour of a

---

<sup>585</sup> (Hayek 1982c, 146)

<sup>586</sup> I add the qualifier "partly" because such fraternity and collective control will be limited to the tasks carried out by the association in question. There will not be a sense that the social order *as a whole* is fraternally subject to collective control.

more collectivist polity. Implicit in Hayek's worry is an assumption about the sources of liberal stability that is quite similar to Rawls: liberal institutions endure only if they are populated by *liberals* who strongly espouse liberal values. We cannot, after all, have liberalism without liberals. This is the clear theme of Hayek's aforementioned essay, "The Intellectuals and Socialism". There Hayek argues that classical liberal institutions will not succeed unless public opinion is brought to support them.<sup>587</sup> Rawls, as we have seen, not dissimilarly thinks that liberal institutions are stable when "those taking part in these arrangements acquire the corresponding sense of justice and desire to do their part in maintaining them".<sup>588</sup> Let us term this the *intentional* account of liberal stability. It matches nicely with the associational view. Liberal institutions endure because the association of liberal citizens who compose them intend this to be the case. There is, however, an alternative account of institutional stability in the liberal tradition more in keeping with the emergent view. Let us simply call this the *emergent* account of liberal stability. This view takes for granted that within the extended liberal order groups will try to organise with the express aim of collectivising society. Still, if the institutional structures in place are effective at inhibiting the formation of such factions and disaggregate political authority sufficiently, the order will endure nonetheless.<sup>589</sup>

Both accounts of stability are long standing within the liberal tradition. The disagreement between John Stuart Mill and Lord Acton over the value of the nation-state was centred on the dispute as to which account is the most appropriate. In his *Considerations on Representative Government*, Mill famously took the view that there is a strong *prime facie* case for ensuring, where feasible, that the boundaries of states and nations coincide,<sup>590</sup> a "nation" being a group of persons united by "common sympathies" that do not extend to outsiders and who desire to be ruled only by themselves or a portion thereof.<sup>591</sup> This fellow-feeling generates the "united public opinion" and shared purposes that Mill thinks is critical to the stability of liberal democratic institutions.<sup>592</sup> For a divided polity is less likely to keep

---

<sup>587</sup> (Hayek 1949, 431–32)

<sup>588</sup> (Rawls 1999, 398)

<sup>589</sup> I refrain here from using the term "modus vivendi," which could be a consciously decided settlement between parties. Emergent conceptions of stability, as I understand them, presume at least the possibility of a disjuncture between the motives of agents and the outcomes. Such stability need not be intended. The emergent conception might best be seen as a *type* of modus vivendi.

<sup>590</sup> Chapter XVI, "Of Nationality, as connected with Representative Government": (Mill 1861, 179–85)

<sup>591</sup> (Mill 1861, 179)

<sup>592</sup> (Mill 1861, 180)

the state within liberal bounds. It is easier under such conditions for the state to adopt divide-and-rule tactics – that “any one of them feels aggrieved by the policy of the common ruler is sufficient to determine another to support that policy”.<sup>593</sup>

Lord Acton criticises Mill in his famous essay “On Nationality”. He too sees co-nationals as united by “a consistent similarity of character, interest, and opinion”.<sup>594</sup> Unlike Mill, he sees the heterogeneous public opinion of multinational states as securing rather than undermining liberal institutions. He envisages each nationality having autonomy as to the laws and norms under which they are governed.<sup>595</sup> Members of such sub-nationalities will develop loyalties towards them since they will constitute “a great source and centre of political sentiments” independent of the central state.<sup>596</sup> The members of each of the different groups will be jealous to protect the autonomy of their grouping. These local loyalties form “a firm barrier against the intrusion of the government beyond the political sphere which is common to all into the social department which escapes legislation and is ruled by spontaneous laws”.<sup>597</sup> None of the autonomous nations *necessarily intend* to keep the state as a whole within liberal bounds when they defend their corporate liberties. Were any of them to have exclusive control of the state, their same political motivations could lead to their imposing their values on members of the other nationalities. However, assuming all goes well such that no one national group can do this, the emergent outcome of all intending to protect their own corporate liberties is a constrained state that cannot dominate any particular group. What enables this outcome is an institutional structure characterised by a diversity of groups and the disaggregation of power.

Along with Mill, Rousseau (if read as a liberal), Locke and Rawls advocate the intentional account of liberal stability while, along with Acton, Hume, Madison and Tocqueville advocate the emergent account.<sup>598</sup> Given that Hayek’s own sympathies lie definitively with the latter set of liberal thinkers, the great emphasis on intentional stability to be found in some of his writings is strange. After all, Hayek stresses above all else that order can result among a large

---

<sup>593</sup> (Mill 1861, 180)

<sup>594</sup> (Acton 1909, 289)

<sup>595</sup> (Acton 1909, 289)

<sup>596</sup> (Acton 1909, 289)

<sup>597</sup> (Acton 1909, 290)

<sup>598</sup> It is best, I think, to see these as ideal-types. Infeasible are both the unanimous consensus on justice of Rawls’ well-ordered society and institutions whose endurance is *solely* contingent on balance of power considerations. The practical question is the *emphasis* of the theorist. I shall return to this point below.



group of parties despite their not intending it. Given also the Hayekian difficulties afflicting coordinated political action, Hayek's own framework gives him good reason to opt for emergent stability. Interestingly enough, when Hayek makes his own institutional proposals, the logic of their operation is most consistent with the emergent conception. We turn to such proposals in the next section.

#### 4. The Feasibility of Emergent Liberal Stability

The aim of this section is to show the feasibility of emergent liberal stability. There are two institutional mechanisms that are conducive to such stability. Institutions might inhibit the formation of partisan interest groups. Institutions marked by a disaggregation of power might also reduce the power that interest groups can acquire and exercise if they are able to form. These strategies are well adumbrated by David Hume, James Madison and Hayek himself. In this section we shall briefly investigate their institutional proposals for how both strategies may be utilised. Note that the claim here is *not* that we should implement one of these institutional blueprints or otherwise a blueprint along the same lines. That would fall foul of the problems with Archimedean institutional reform that we identified in chapter two.<sup>599</sup> Instead we study these institutional proposals to try and show that classical liberal institutions utilising emergent stabilising forces are perfectly conceivable. We shall return to this point later in this section.

David Hume's thinking about institutional stability approaches the problem by considering how the problem of faction in politics can be dealt with. His essay "of parties in general" provides a useful typology.<sup>600</sup> Factions can be *personal* in that they are not based on any substantive sentiment or interest. For the human predisposition to form competing groups, Hume observes, can be triggered by even the most trivial of differences. They can also be based on material or economic *interest*, subscription to an abstract *principle*, or the *affection* of a charismatic leader or family. These are ideal-types with actual factions usually being "mixed".<sup>601</sup> Hume is concerned about such groups for the same reason as is Hayek. Factional

---

<sup>599</sup> Even if comprehensive design of institutions were feasible, it would not be advisable to simply follow a blueprint. For the most appropriate institutional structure would probably vary from context-to-context. There are, after all, many ways to ensure that authority is disaggregated.

<sup>600</sup> (Hume 1777a, 59–63)

<sup>601</sup> He describes 18<sup>th</sup> century Britain's political parties as such - Hume (1777b).

groups are dangerous because they can use the political machinery of the state to further their own interest at public expense. Hume's institutional proposal, the "Idea of the Perfect Commonwealth," is advanced primarily to show how the problem of faction can be mitigated. There he argues against the assumption made by Montesquieu and Rousseau that faction can only be avoided in small, homogenous states marked by little variation in interest or sentiment.<sup>602</sup> Hume turns on its head the logic of this position. In a small state, the proximity or "near habitation" of the parties makes it relatively easy for factions to form and capture political power. Instead we should opt for larger states in which factions are more difficult to form. In such states "the parts are so distant and remote, that it is very difficult, either by intrigue, prejudice, or passion, to hurry them into any measures against the public".<sup>603</sup> Thus, Hume says that it is "probable" that the perfect commonwealth would be realised in a large rather than small state. The substance of the rules regulating the perfect commonwealth all push towards the radical decentralisation of governmental authority to prevent the concentration of power in any one part of the state.<sup>604</sup>

His essay would be a substantial influence on James Madison's contributions to *The Federalist*.<sup>605</sup> Madison understands a faction similarly to Hume, as a subset of citizens "united and actuated by some common impulse of passion, or of interest, adverse to the rights of other citizens, or to the permanent and aggregate interests of the community".<sup>606</sup> He too thinks that a large state constitutes an institutional setting more difficult for the formation of faction. For when you:

Extend the sphere [of the state], and you take in a greater variety of parties and interests; you make it less probable that a majority of the whole will have a common motive to invade the rights of other citizens; or if such a common motive exists, it will be more difficult for all who feel it to discover their own strength, and to act in unison with each other. Besides other impediments, it may be remarked that, where there is a consciousness of unjust or dishonorable purposes, communication is

---

<sup>602</sup> (Hume 1777c, 528–30)

<sup>603</sup> (Hume 1777c, 230)

<sup>604</sup> He proposed to divide the United Kingdom into 100 counties each of which in turn is divided into 100 parishes. The voters in each parish, restricted to property holders, would elect a representative. The hundred representatives of each county would elect out of their number a senator and ten magistrates. There are, then, two layers of representation between those charged with governance and the actual voter. Executive power would lie with the senators and legislative with the magistrates. Further, the magistrates would never meet as a whole body but would convene in their own counties. (Hume 1777c, 516–620)

<sup>605</sup> As convincingly argued by Adair (1957).

<sup>606</sup> (Madison 2003a, 41)

always checked by distrust in proportion to the number whose concurrence is necessary.<sup>607</sup>

In a large state with a diverse population and diverse interests, the formation of faction is less likely. Even where there is a common motive, coordination problems constitute an extra impediment to the formation of faction. Madison too and quite famously favours the disaggregation of political authority that hinders factions when they do manage to form. He proposes to “divide the legislature into different branches; and to render them, by different modes of election and different principles of action, as little connected with each other as the nature of their common functions and their common dependence on the society will admit”.<sup>608</sup>

Hayek himself utilises both strategies in his own institutional proposals. The first of inhibiting the formation of faction is on display in his essay, “The Economic Conditions of Interstate Federalism,” first published in 1939 immediately before the Second World War. Hayek there argues that a political and economic union of European countries is required if peace is to be secured in the future. For, “[t]he existence of any measure of economic seclusion or isolation on the part of an individual state produces a solidarity of interests among all its inhabitants and conflicts between their interests and those of the inhabitants of other states”.<sup>609</sup> At the limit, this conflict of interest leads to war. Hayek here makes the empirical assumption that a regime of economic protectionism encourages the concentration of critical industries or sectors in particular states. The economic interests of the whole nation will be tied up with these critical sectors. In consequence, “any change which affects a particular industry in a certain territory should impinge more heavily upon all or most of the inhabitants of that territory than upon people elsewhere”.<sup>610</sup> He continues, “[t]he mere fact that everybody will find again and again that their interests are closely bound up with those of one constant group of people and antagonistic to that of another group is bound to set up severe frictions between the groups as such”.<sup>611</sup>

An economic union consisting of the free movement of people and capital can help to alleviate this problem. Under such circumstances the composition of groups will be less permanent. Hayek’s implicit but unstated assumption here is that a freer economic regime

---

<sup>607</sup> (Madison 2003a, 45)

<sup>608</sup> (Madison 2003b, 253)

<sup>609</sup> (Hayek 1948, 257)

<sup>610</sup> (Hayek 1948, 257)

<sup>611</sup> (Hayek 1948, 258)

will be more dynamic and subject to Schumpeterian creative destruction. Sectoral change will be more common and interest groups will have less time to solidify.<sup>612</sup> Additionally, under such conditions it is less likely that the boundaries of key sectors will overlap with territorial boundaries. The conflicts, then, would not be between the same groups but of “groups of constantly varying composition... the various individuals finding themselves arrayed, sometimes with one group of people against another, and at other times on another issue with the second group against the first”.<sup>613</sup> The political impact of economic union is to make it much more difficult, then, to form *durable* factions. Additionally, the diversity of national groups in such a union would make it much more difficult to create political coalitions for central planning. In national states, on the other hand, it is easy for an interest group to “persuade the rest of the community that it is in their interest to protect ‘their’ iron industry or ‘their’ wheat production or whatever it be”.<sup>614</sup> Tellingly for our purposes, here Hayek cites positively Acton’s critique of Mill’s nationalism.<sup>615</sup>

The chapter on Hayek’s model constitution in the third volume of *Law, Legislation & Liberty* begins with an epigraph from Hume’s “Idea of the Perfect Commonwealth”. It also states its aim to be the same as that of the founders of the US Constitution, though Hayek thinks with the benefit of hindsight one might be able to improve on their efforts.<sup>616</sup> If in the interstate federalism essay he employs the strategy of inhibiting the formation of faction, the emphasis here is on the disaggregation of authority. His aforementioned concern in this work is the legislative capture of democratic institutions by coalitions of special interests. Hayek’s proposes an institutional framework in which the disaggregation of authority extends further than those found in actually existing liberal democracies. In essence, Hayek proposes to divide our current legislative chambers into two bodies. The first is called the “legislative assembly”. Their function is to produce the abstract rules of conduct on which the extended order depends. There are two aspects to this task. First, to formulate abstract rights of person and property that apply to all. These form the common rules which structure interaction in the extended liberal order.

---

<sup>612</sup> C.F. (North, Wallis, and Weingast 2009, 25): “Creative economic destruction produces a constantly shifting distribution of economic interests, making it difficult for political officials to solidify their advantage through rent-creation”.

<sup>613</sup> (Hayek 1948, 257)

<sup>614</sup> (Hayek 1948, 262)

<sup>615</sup> (Hayek 1948, 270)

<sup>616</sup> (Hayek 1982c, 105)

Second, they formulate constraints on the manner in which the state provides services to citizens.<sup>617</sup> The legislative assembly does not so much legislate policies itself but sets what we might consider as “Hayekian constraints” on policy-making. In Hayek’s estimation they require a particular kind of representative. His somewhat curious proposal is that the candidates would need be 45 years old and “have already proved themselves in the ordinary business of life,” and they would be elected by peers of the same age, governing for 15 year terms.<sup>618</sup> The rationale here is that their role requires politicians with a long time horizon. The second chamber, the “governmental assembly” will resemble the legislative chambers with which we are familiar. They would be staffed by conventional politicians who sought electoral majorities in regular elections. Though, Hayek thinks that such assemblies can be more numerous, governing much smaller jurisdictions than their contemporary equivalents.<sup>619</sup> They would also legislate under much greater constraints. Hayek hopes that this makes them much less vulnerable to electoral capture by interest groups. Importantly, in this essay he re-affirms his earlier commitment to an international federation.<sup>620</sup>

Now, as already mentioned above, the purpose here is not to advocate the top-down implementation of any of these institutional proposals. This thesis is sceptical as to the possibility of such reform. Furthermore, one can certainly identify difficulties with Hayek’s proposals. The concern with the international federation is that any central authority strong enough to enforce the rules of the economic union might be liable to abuse its power in other ways. The concern with his model constitution is that while it reduces the power of governmental assemblies, it does so by greatly empowering the legislative assembly.<sup>621</sup> If we are worried about the undue sway of interest groups in democratic life perhaps super-majority voting rules, as proposed by James Buchanan and Gordon Tullock, would be a simpler and more effective institutional mechanism.<sup>622</sup> Given these specific concerns about Hayek’s institutional proposals and the thesis’ wider scepticism as to our ability to implement such institutional blueprints, one might wonder what function they perform in the argument of this chapter.

---

<sup>617</sup> (Hayek 1982c, 445)

<sup>618</sup> (Hayek 1982c, 449)

<sup>619</sup> (Hayek 1982c, 479)

<sup>620</sup> (Hayek 1982c, 482)

<sup>621</sup> On this see Müller (2015).

<sup>622</sup> (Buchanan and Tullock 1999); c.f. (Gaus 2012a, 540–45)

I suggest we think of these institutional proposals not as action-guiding prescriptions but rather as theoretical models illustrating that emergent liberal stability is feasible. Recall that Hayek's worry is not that we cannot engineer liberal institutions where they do not exist but that they will have a limited propensity to endure *where they do arise* because they will not win the loyalty of their subjects. We have shown, drawing in part on his own institutional models, that this is not necessarily the case. Conceivable are liberal institutions that lack such loyalty but endure because they effectively disaggregate power and inhibit the formation of faction. Still, it might be objected that models conjecturing how emergent stabilising forces *might* be feasible do not show that they indeed *would* be feasible. In response, we can observe that the institutional models of Madison, Hume and Hayek are not mere theoretical conjecture but reflect historical constitutional practice. For instance, on one interpretation, liberal institutions and practices arose and were sustained in Europe as a result of competition between competing and overlapping jurisdictions.<sup>623</sup> In other words, their endurance was secured by emergent stabilising forces.

Another concern might be that it is not particularly encouraging for those who share Hayek's commitments to know that liberal institutions utilising emergent stabilising forces might endure where they do arise but, per the contention of this thesis, there is little we can consciously do to create them where they do not. This point must be conceded to as an implication of this thesis, which indeed suggests that societies stumble upon liberal institutions largely as a matter of contingent, historical accident. In some societies the interactions between various actors and agencies will lead to convergence on such institutions but in others they will not. North, Wallis and Weingast's argument, which we first saw in chapter three, is precisely that the checks-and-balances that we consider part-and-parcel of Western constitutional practice arose as the unintended consequence of strategic interactions between political elites over successive generations. Note though that this point cuts both ways. If liberal institutions can't be consciously created when they are *not* in place, then by the same token they are difficult to upend when they *are* in place. It might also be the case that the elaborate institutional proposals advocated by Madison, Hume and Hayek are not a necessary condition to achieving at least the disaggregation of authority. The latter might simply require a generalised norm of toleration towards dissent and refusing to try and centrally regulate the social order. Perhaps under circumstances where that norm was

---

<sup>623</sup> On this see Salter and Young (2019). Another relevant historical example is the Ottoman Empire, which we shall discuss briefly in the section to follow.

upheld, we might see the proliferation of different institutional authorities.<sup>624</sup> These numerous associations would emerge in a spontaneous fashion and they would also check one another's ambitions in a spontaneous fashion. They would thus be emergent in both senses in which the term has been used in this thesis. A final and important worry, articulated by John Rawls among others, is that emergent stabilising forces are insufficient to engender the long-run stability of liberal institutions. The next section will deal with this important objection at length.

## 5. The Rawlsian Objection to Emergent Stability

The Rawlsian objection to emergent stabilising forces runs as follows. Institutions that rely upon such a stabilising force will not long endure because they are at the mercy of changes in the distribution of power. Since none of the factions have any intrinsic commitment to liberal institutions, the institutional constraints will be upended as soon as a faction or coalition of factions have the political power to do so. The mere *modus vivendi* of the emergent conception is insufficient to achieve a sufficient degree of institutional stability. As Jacob Levy articulates this Rawlsian view: “[t]here is no constitutional design so clever, no multiplication of veto points so elaborate that the liberal order can be formally guaranteed. There need to be some actual liberals involved”.<sup>625</sup> In other words, we can't have a liberalism without genuine liberals because institutional constraints on illiberal political ambitions can bear only so much of that weight. Now, we have already argued that classical liberal institutions might well produce subjects who endorse the manner of their operation. Still, we are setting that argument aside here and trying to respond to this objection by defending the emergent conception of liberal stability on its own merits.

There are three responses we can make. First, there are historical precedents of comparatively decent regimes that achieved accommodation between different groups without being subject to deep endorsement by any of them. If that is true of such regimes, it should be true of liberal institutions reliant upon emergent stability. Take the Ottoman Empire. Its system of *millets* protected the rights of non-Muslim minorities to practice their religion and live according to their own customs. Bernard Lewis writes:

---

<sup>624</sup> I here follow Chandran Kukathas (2003) model of a “liberal archipelago”. C.f. Pennington (2015)

<sup>625</sup> (Levy 2018, 155).

A *millet* was a religio-political community defined by its adherents to a religion. Its members were subject to the rules and even to the laws of that religion, administered by its own chiefs. . . In return for this measure of religious freedom and communal autonomy, non-Muslim *millets* owed allegiance to the state.<sup>626</sup>

Given their subservient status, it seems plausible to assume that the minorities within it did not endorse the regime in any deep or substantive sense. Indeed, the eagerness with which many broke away from the Ottoman regime as it disintegrated in the early twentieth century suggest that they did not. Still, it was a comparatively decent regime, particularly when compared to its European neighbours at the time, and it lasted for several centuries. No actually existing liberal democracy has endured for such a length of time. This suggests that it is perfectly possible for regimes that lack the deep endorsement of the governed to endure over the long run.

This leads us to our second response. The Rawlsian approach to thinking about liberal stability amounts to something of a false dilemma. *Either* all or nearly of the parties deeply endorse the institutions and are fully motivated to comply *or* the parties are purely strategic actors who value institutions only instrumentally and are prepared to defect as soon as the cost-benefit calculus is in their favour. The error with this view is that it overlooks more ambiguous political motivations that lie in between these two extremes. If we reject deep Rawlsian consensus, then we must be positing a world of purely Hobbesian instrumentalists. However, even if I live under a political regime whose tenets are out of sync with my deepest motivations, it does not follow that I will behave only as a knave. I might still place at least some intrinsic value on complying with certain conventions in political life. There is probably no sphere of life in which interactions take place on purely strategic terms. Even power relations between states, from which the term “modus vivendi” was originally derived, are not motivated solely by strategic self-interest.

Such a perspective is gleaned from Hedley Bull’s *The Anarchical Society*, a key text in the “English” school of International Relations.<sup>627</sup> According to his theory, in some contexts the international relations between states can constitute an *international society*.<sup>628</sup> Such a society exists when the states in question “conscious of certain common interests and common values. . . regard themselves as bound by certain rules in their dealings with one

---

<sup>626</sup> Cited in Gray (2000, 152).

<sup>627</sup> In contemporary international theory, the constructivist school takes a similar view to Bull. See Wendt (1992).

<sup>628</sup> I draw here primarily from Bull (2012, 13–18)



another".<sup>629</sup> These rules might include respecting one another's claims to independence, honouring agreements and exercising force on one another subject to certain limitations. The critical point is that the political elites did not view these rules entirely strategically but placed some measure of intrinsic value on complying with them. The relations between post-Westphalian European states or Ancient Greek city states constitute international societies on Bull's score. He certainly does not exaggerate the intrinsic commitment to these rules. They were applied often partially and selectively. In the European case, the interests of small states in particular were overlooked as they were progressively absorbed by their larger neighbours between 1648 and 1848. However, their presence can be verified by the different form that *intra*- and *inter*-society relations respectively took. Thus, while European states recognised at least a putative norm that they should fulfil their promises to one another, there was no expectation that agreements entered into with the Ottoman Empire were binding. Similarly, in their relations with one another Greek city states recognised certain constraints on how war should be conducted. They did not recognise such constraints in their relations with the Persian Empire.<sup>630</sup> I take Bull's point to be that if states' foreign policy was dictated purely by strategic considerations, we should not see this kind of difference.

If power relations between states can be mediated to some degree by intrinsic commitment to the rules, then perhaps something similar can be said about the domestic sphere. This might explain the observation of many realist political theorists that actually-existing liberal democracies seem to cope reasonably well with citizens whose political motivations are much more contingent and apathetic than their hypothetical counterparts in Rawls' well-ordered society.<sup>631</sup> Despite their often less-than-full endorsement of their societies, these citizens are not complete knaves. I believe the following thought experiment can exemplify this point. Imagine, say, a conservative Christian party. They take the view that traditional Christian morality with respect to matters such as homosexuality or abortion or blasphemy should be enforced in public life. We might say that their first-best regime is one where society is united in upholding Christian values and goals. Suppose they are allowed to freely operate in a liberal political society that nonetheless contains an elaborate system of checks

---

<sup>629</sup> (Bull 2012, 13)

<sup>630</sup> The manner in which Europeans would treat non-Europeans during colonialism is also surely relevant here.

<sup>631</sup> E.g. (Gray 2000); (McCabe 2010); (Horton 2003)

and balances to ensure that no group, including them, can comprehensively succeed in their legislative agenda. Suppose also that the system is performing reasonably as intended. Our Christian party might gain *some* victories here and there. A tax-break for marriage perhaps or some moderate regulation of abortion but overall the results of their efforts are quite limited. This is so because they are neutered, in effect, by factions with opposing views and other constitutional constraints built into the political institutions. In turn, our conservative Christian party acts as a counterbalance to the legislative agenda of many of their rivals.

The following I believe is a plausible picture of the party's motivational attitudes. The constitutional framework is certainly not their first-best. Were the opportunity available to campaign for constitutional change, they would campaign for a system which would make it easier for them to win legislative victories and correspondingly harder for their political opponents. They look forward to the day when enough of their fellow citizens come round to their way of thinking making such a constitutional change viable. Indeed, they are happy to "play politics" with certain constitutional procedures to rig the rules in their favour. If they are to decide or to have a say in deciding how, say, constituency boundaries are to be reformed, they will try to push for reforms that benefit them electorally. If, likewise, they have a say in selecting judicial officials, they will try to pick those that are sympathetic to their view.

Still, this does not mean they take a *purely* strategic or instrumental view of constitutional procedures. That the constitutional conventions under which they politically labour are a time-honoured practice might engender a certain kind of fidelity – they might think: "for all its faults, that this is how we doings around here makes it worthy of some respect". Relatedly, a history of contestation and sometimes collaboration with their political opponents could generate some degree of good-will towards them. The upshot is that their strategic motivations are moderated by some sense of "fair play" even when operating under political institutions which are inconsistent with their deepest commitments. Thus they might affirm that one does not disenfranchise opposing party voters or refuse to acknowledge a clear electoral victory of one's opponent or tamper with the counting of ballots. They would maintain their commitment to these norms even when a purely strategic calculus might make breaking them seem worthwhile. Still, this falls some way short of the deep commitment that Rawlsians hope all citizens would have towards liberal institutions.

The case of this political party is entirely a thought experiment. I hope its force lies in its familiarity. I think many political parties and factions in actually-existing liberal democracies adopt this complex and ambiguous attitude towards the political institutions under which they operate. This seems especially true of political participation by groups who by their nature cannot deeply endorse their political institutions, such as secessionist groups.<sup>632</sup> Actually-existing liberal democracies already appear reasonably stable despite their frequent reliance on what we might term these “ambiguous loyalties”. The stability of a classical liberal order might likewise be predicated in part on these ambiguous loyalties and also in part on institutional features that disaggregate power and hinder the formation of faction. It does seem plausible to say that not even the most byzantine checks-and-balances can ensure stability if vast swathes of persons actively want to undermine existing institutions. Neither, however, do we need posit that they all be fully committed liberals either, of the Hayekian or Rawlsian or any other variety.

The third response is to suggest that the Rawlsian objection in any case under-estimates the efficacy of strategic or self-interested motivation alone in securing stable and long-term political cooperation. Here we can draw on conceptual and sociological claims made by Russell Hardin. Before investigating Hardin’s arguments we should revisit in some detail why Rawls thinks deep endorsement of liberal institutions is crucial to their endurance.<sup>633</sup> As we saw in chapter four, Rawls’ assumes that once liberal citizens are committed to liberal values, all that is required for stable political cooperation is that they assure one another of this fact.<sup>634</sup> In the language of game theory, they play an assurance game. The pay-offs for such a game were displayed in figure 4.1 in chapter four. However, he also thinks that if liberal citizens lack an adequate sense of justice – i.e. if they are solely self-interested actors – then they are engaged not in an assurance game but a prisoner’s dilemma, the pay-offs to which are showed in figure 7.1:

---

<sup>632</sup> From my anecdotal observations, the Scottish National Party, who are committed to Scottish independence, do not seem less committed to British parliamentary principles and practices than other parties, despite their evident lack of deep loyalty to the British state.

<sup>633</sup> As in our discussion of the Rawlsian assurance game in chapter four, the interpretation of Rawls to follow draws greatly on Paul Weithman’s (2011) reading, particularly chapter two of that work.

<sup>634</sup> Though we also observed that this is much more easily said than done.

		<b>Player 2</b>	
		<b>Comply</b>	<b>Defect</b>
<b>Player 1</b>	<b>Comply</b>	2 <sup>nd</sup> 2 <sup>nd</sup>	1 <sup>st</sup> 4 <sup>th</sup>
	<b>Defect</b>	4 <sup>th</sup> 1 <sup>st</sup>	3 <sup>rd</sup> 3 <sup>rd</sup>

**Figure 7.1 – The Prisoner’s Dilemma**

As with figure 4.1 from chapter four, the pay-offs are represented in an ordinal fashion with the 1<sup>st</sup> being the highest and the 4<sup>th</sup> being the lowest. The difficulty here is that cooperative relations are much more difficult between players caught in a prisoner’s dilemma. Recall that in the assurance game none of the players have a dominant strategy worth executing regardless of what the other player does. One is better off complying if the other player does and vice-versa better off defecting if the other player does. Matters differ in the prisoner’s dilemma. The pay-offs for each are such that they are better off defecting no matter what the other counter-party does - (1<sup>st</sup> > 2<sup>nd</sup>) if they comply and (3<sup>rd</sup> > 4<sup>th</sup>) if they defect. Both parties will defect. This is the “Nash equilibrium”. If both are defecting neither can improve their position by switching to a compliance strategy. Each will gain a higher pay-off if both complied but joint-compliance is not an equilibrium strategy. For if both are complying, each can improve their position by defecting. Self-interested behaviour leads to sub-optimal outcomes for all. One solution to the problem is to utilise external sticks and carrots to change the incentives facing our players to make it worth their while complying. This solution Rawls associates with Hobbes.<sup>635</sup> The problem with this solution is that the self-interested parties to a prisoner’s dilemma are fundamentally not trustworthy. Players receive the highest possible pay-off when the others comply and they defect i.e. each is more than happy to free-ride on the efforts of others. While the hope would be to utilise a system of external sanctions that would minimise if not eliminate opportunities to free-ride, opportunistic actors will always find *some* loop holes or otherwise renege as soon as the balance of power is in their favour. Once one player defects, it is in the interests of the other to do so also and

---

<sup>635</sup> (Rawls 1999, 238, 304–5)

cooperation will break down.<sup>636</sup> It is for this reason that Rawls thinks institutional constraints alone are insufficient to maintain cooperation.

The only way we can reliably secure compliance by the parties is to change their *internal* motivation structure so they place intrinsic value on compliance. Once parties acquire an intrinsic commitment to liberal justice, their pay-offs change such that they are no longer playing a prisoner’s dilemma but an assurance game as displayed originally in figure 4.1. The pay-off table for the latter game is reproduced below in figure 7.2:

		<b>Player 2</b>	
		<b>Comply</b>	<b>Defect</b>
<b>Player 1</b>	<b>Comply</b>	1 <sup>st</sup> 1 <sup>st</sup>	2 <sup>nd</sup> 4 <sup>th</sup>
	<b>Defect</b>	4 <sup>th</sup> 2 <sup>nd</sup>	3 <sup>rd</sup> 3 <sup>rd</sup>

**Figure 7.2 – Mutual Assurance Game**

The key difference is that since there is no dominant strategy, there is no single Nash equilibrium in the mutual assurance game. Joint-defection is an equilibrium since no party can improve their pay-off by changing strategy under such circumstances. Joint-compliance, however, is also a possible equilibrium. Further, each gain their highest pay-off if both comply. So long as players can reliably signal to one another their commitment to comply, they can arrive at and stay in a stable equilibrium of joint-compliance.<sup>637</sup> Cooperative behaviour in political life is thus much more viable if the agents are playing an assurance game. We do not need to worry about their defecting as soon as a loophole presents itself. For the players themselves have no wish to defect from the stable cooperative equilibrium. For Rawls a widespread sense of justice turns political life from an intractable prisoner’s dilemma into a resolvable assurance game by giving liberal citizens intrinsic motivation to comply with justice so long as others reciprocate.

---

<sup>636</sup> Chapter 4 of *Political Liberalism*, “The Idea of Overlapping Consensus” articulates Rawls’ concerns here at length. The “social unity” between self-interested actors, he writes there, “is only apparent, as its stability is contingent on circumstances remaining such as not to upset the fortunate convergence of interests”. (Rawls 1996, chp. 4, 147)

<sup>637</sup> Of course, as we saw in chapter four, acquiring this assurance in a large-scale society, particularly as it relates to the internal motivations of other actors, is not as straightforward as Rawls assumes.

The Rawlsian objection to the utilisation of emergent stability is based, then, upon the presumption that self-interested actors would face the pay-offs of a prisoner's dilemma. Russell Hardin, however, disputes this assumption.<sup>638</sup> On Hardin's Humean view, the Rawlsian assumption that strategically motivated actors would play a prisoner's dilemma in the absence of a strongly internalised sense of liberal justice is mistaken. For Hardin, Rawls and other contractualists make this mistaken assumption because they think participation in political life is akin to a commercial contract. Viewed in this way, political interaction looks like a prisoner's dilemma because contractual relations often provide opportunities for the parties to renege and improve their position. However, Hardin observes, political interactions is decidedly not like this:

One can renege on any given contract and plausibly still keep open the opportunity for mutually beneficial contractual relations with other potential partners. But one cannot will away the whole institution of enforcing contracts and then still expect mutually beneficial contractual exchanges with anyone to work.<sup>639</sup>

Precisely because it is in no one's interest to "will away the whole institution of enforcing contracts," Hardin argues that strategically motivated actors, pace Rawls, are *already* playing an assurance game.<sup>640</sup> Positing a strongly internalised sense of justice is unnecessary. Hardin also thinks that the pay-off for the first-best outcome *far exceeds* that of the second-best outcomes and the difference between the second- and the third- and fourth-best outcomes respectively are trivial in comparison.<sup>641</sup> Thus, if our strategic actors have some reasonable assurance that others will comply, they will have overwhelmingly *self-interested* reasons to comply themselves, unlike in the prisoner's dilemma scenario. Hence, political constitutions, he writes, are "self-enforcing".<sup>642</sup> Why might actors have an overwhelming interest to comply with the "whole institution of enforcing contracts" and not renege from it as they might do when engaged in the everyday transactions made possible by such an institution? Hardin provides two answers to this question. The first is that everyone is made better off under such an institutional framework which moves us away from a subsistence economy into a productive one.<sup>643</sup> This strikes me as only partially true given that states have often

---

<sup>638</sup> I here draw on chapter 4 of his *Liberalism, Constitutionalism and Democracy* (Hardin 1999).

<sup>639</sup> (Hardin 1999, 87)

<sup>640</sup> (Hardin 1999, 92-98)

<sup>641</sup> (Hardin 1999, 92-98)

<sup>642</sup> (Hardin 1999, 92-98)

<sup>643</sup> (Hardin 1999, 98)

incorporated individuals and communities into market institutions regardless of their own wishes simply because they are subsequently easier to regulate and tax.<sup>644</sup>

Another reason Hardin offers is that once these institutions gain a critical mass of followers network affects are generated which markedly changes the cost-benefit calculus of those who are yet to sign unto them.<sup>645</sup> He provides a case study of such network affects in action: the ratification of the US Constitution.<sup>646</sup> Its contents, particularly the Commerce Clause, were framed for the benefit of the commercial class, namely large business and plantation owners. At the time of its formulation, the Constitution would have been opposed by the bulk of small farmers who composed the large majority of US property owners. However, the ratification process was manipulated such that the Constitution was put to vote first in those states in which it faced little or no opposition. Once these states ratified, other states who may initially have been opposed now felt that they had strong self-interested reasons to comply:

By staging votes in the states where there was little or no opposition first, the Federalists were able to make the votes of later states a matter of joining the coordination or accepting a world even worse than that of the Articles of Confederation, a world in which they would be outside the nation and unable to enjoy the benefits it offered and, just possibly, at risk of war from the more populous new nation.<sup>647</sup>

In other words, as each new agent endorses the constitution, so increases the costs of holding out for the remaining players, giving them strong reasons to comply themselves. In the North American case, the hold-outs would eventually ratify not simply to avoid the risk of war but also to avoid the economic isolation that comes with being on the other side of an external tariff barrier. In many ways, this is a restatement of the point that we already made earlier in the thesis when considering Gaus' Kantian coordination game or Mainsbridge et al's concept of an emergent "societal decision". Institutional rules can emerge and become entrenched without the deep endorsement of those subject to their authority. This, I think, is normatively worrying in many ways. But it also shows that prudential considerations can go much further in securing political and social cooperation than Rawls allows.

---

<sup>644</sup> On this see Scott (1998, Chp. 1).

<sup>645</sup> Though the use of "network affects" is my terminology rather than Hardin's.

<sup>646</sup> (Hardin 1999, 103–7)

## **Conclusion**

We have tried to alleviate Hayek's worry about the viability of classical liberal institutions on two fronts. The stability of such institutions can be secured either through the endogenous socialisation of those subject to their authority or through emergent stabilising forces that hinder the formation of faction and reduce its potency when it can form by disaggregating political authority. In practice, the stability of liberal institutions can be secured through a mixture of both these mechanisms. The final section responded to the Rawlsian objection that emergent stabilising forces will falter unless liberal citizens deeply internalise liberal values. The manner in which we have tried to understand liberal stability in this chapter is in keeping with the general tenor of this thesis. Large-scale outcomes, including even the endurance of liberal political institutions, are emergent phenomena outrunning the intentions of the actors involved. For most of us, most of the time, politics is experienced as something that happens to us rather than something we control. We find ourselves in a political and social context and try to make the best of it. The concluding section will make some remarks as to which understanding of freedom might be most helpful in this regard.



## Thesis Conclusion: The Freedom “Inherent” in Agency

This thesis has argued against collective political agency. The theme running throughout has been that mutual ignorance prevents impersonally related citizens from forming and acting upon shared intentions as to how their institutions should be structured. It is not simply that we do not live in a political world of our own making, there really is no meaningful referent to the “we” that such a claim deploys. The two great goods promised by the associational understanding of political society, collective control and social unity, are not within our grasp according to the account that has been advanced and defended here. This is the price of living in the impersonal liberal order in which we find ourselves. Hayek thought that for many of us the cost is too much to bear. There will be a perennial attraction to political projects that try to reconcile our desire for freedom with our desire for unity and control over our social world. The last chapter alleviated Hayek’s worry that this would undermine the stability of liberal institutions. This conclusion offers some reflections on what kind of orientation in the world might help us make peace with the liberal order. Particularly, the aim is to offer some remarks as to what understanding of freedom can help us do so.

Undoubtedly, if one thinks about freedom in a Rawlsian fashion, the arguments of this thesis are most dispiriting. Taking its debt from Rousseau as much as Kant, the Rawlsian understanding of freedom is in an important sense deeply collectivist. On this view, I do not attain my full autonomy unless the basic structure of my society accords with a conception of justice that I affirm along with all my other fellow citizens. In other words, I am free only if I live in a social world over which the citizen-body, of which I am but one part, exercises collective control. If the arguments of this thesis are correct, then under the impersonal conditions of modernity, we cannot be free. Undoubtedly this seems like an unhappy result. Fortunately, there is another way of thinking about freedom under modern conditions. This understanding of freedom is to be found in the work of Michael Oakeshott.

As a segue way into our discussion of Oakeshott, it would prove useful to reflect on why the exercise of freedom without collective control and social unity seems inadequate. Such freedom might lack *meaning*. This is the problem that Hegel identified with the bare subjective freedom available to us as isolated members of civil society. I am but one of millions of citizens who compose a nation-state. On a more Hayekian construal, we might say that I am but one of billions of persons who compose the extended order of cooperation, or “Great Society,” which by now covers the globe. I am also but one among countless human

beings who will inhabit this planet in the lifetime of the human species. These facts alone might be enough to provoke the existential worry that one's singular life is irrelevant. If we think in even grander terms, of our place in the wider scheme of the universe, our singular lives seem to occupy a category beyond irrelevance. All this, Thomas Nagel once famously observed, lends the human conditions its absurdity. The seriousness with which we pursue our lives when we are focused on living them seems ridiculous when, from a more Archimedean perspective, we dwell on how irrelevant if not futile our efforts are in the grand scheme of things.<sup>648</sup> This absurdity is not always easy to bear. We wish that somehow our individual exercise of freedom connects to something of greater significance. Contextualising our own life story as at least partly a contribution to a grand scheme of cooperation in pursuit of shared goals, from one generation to the next, is one way of giving it this wider significance. If the value of our freedom is contingent on its being meaningful in *this* sense, then the arguments of this thesis suggest that such meaning is not to be found.

Freedom without collective control and social unity might also be *insecure*. If the social world which I inhabit is subject to unpredictable and uncontrollable changes, I will not always be able to secure my goals. For many, the value of freedom is questionable unless they can have some confidence, if not a guarantee, that the objectives they use their liberty to pursue can indeed be obtained. Concerns about social justice can often be interpreted as a concern about the security of our freedom. It is unfair that some—the wealthy or the privileged—have access to opportunities that the less fortunate do not. They are better placed to secure the objects of their freedom than are the less advantaged. Above all, the intuition driving the Rawlsian idea of a just society is that the freedom of all liberal citizens should be made as secure as possible. One way of increasing the security of one's freedom is by changing the wider institutional framework – the basic structure in Rawls' terms – within which we live. The lone person can exercise agency only *within* this institutional structure but not *over* the structure itself. This is something we can only do together as a citizen body. In Margaret Moore's terms, citizens engaged in political participation experience a "different (collective) dimension of autonomy". Chapters two and three in particular criticised this notion in various ways. There is no Archimedean point from which major institutional rules, understood as a holistic basic structure, can be reformed at will. States are not the intentional projects of a civic minded citizenry but emergent phenomena that arise from human action and yet not by design. (Though the "actions" of the politically well-connected may have much more sway

---

<sup>648</sup> (Nagel 1971)

over events.) The capacity of the citizenry to ensure that government operates in line with their views and preferences is much exaggerated. If they find themselves in an institutional setting where their freedom is secure in the sense that concerns Rawls and other liberal egalitarians, they are above all *lucky*.<sup>649</sup> If they don't find themselves under such institutional settings, they are not going to change it through the force of an intentional will they share with their peers.<sup>650</sup>

Both these fears arise to the extent that we value freedom instrumentally. If the value of freedom is to be found in the objects it allows us to pursue, then freedom is as valuable as our expectation of securing the objects of our pursuit. The insecurity of that pursuit diminishes the value of freedom. If the value of freedom is to be found in what it helps us achieve, then we worry about the lasting significance of that achievement. Reflection on the wider insignificance of our achievements diminishes the value of our freedom. Michael Oakeshott's reflections on the nature of freedom in his work, *On Human Conduct*, point towards a different way of thinking about freedom. Oakeshott is keenly aware that accompanying freedom is the risk of both failure and meaninglessness – "The sort of self-fulfilment it promises is partnered by the notorious risk of self-estrangement and self-destruction".<sup>651</sup> Freedom so understood can feel like a heavy burden. Certainly, historically, for many the "opportunities of self-disclosure and self-enactment which emerged from the desuetude of a communally organized life" did not seem worthwhile.<sup>652</sup> Oakeshott also reflects on the kind of personality who, by contrast, adopted a disposition that helped them to make something of these opportunities. Such a disposition:

... displayed itself in the persons of the younger sons making their own way in a world which had little for them, of foot-loose adventurers who left the land to take to trade, of town-dwellers who had emancipated themselves from the communal ties of the countryside, of vagabond scholars, in the speculative audacities of Abelard, in venturesome heresy, in the lives of intrepid boys and men who left home to seek their fortunes each intent upon living a life for "a man like me", and in the relationships of men and women.<sup>653</sup>

What distinguished this kind of personality for Oakeshott? A disposition to cultivate the freedom "inherent" in agency. The freedom to choose should be made to "yield a satisfaction

---

<sup>649</sup> Danish citizens did not create Denmark. They were fortunate enough to be born there.

<sup>650</sup> One of the great and I think under-appreciated Hayekian lessons of North, Wallis and Weingast's *Violence and Social Orders* is that the open-access liberal order with which we are familiar were emergent phenomena. No one willed it. This was the key claim advanced in chapters two and three.

<sup>651</sup> (Oakeshott 2003, 236)

<sup>652</sup> (Oakeshott 2003, 275)

<sup>653</sup> (Oakeshott 2003, 239)

of its own, independent of the chancy and intermittent satisfaction of chosen actions achieving their imagined and wished-for outcomes". The process of "imagining, deliberating, wanting, choosing and acting" should not be seen as costs to be incurred in the pursuit of other ends but the "exercise of a gratifying self-determination or personal autonomy".<sup>654</sup>

Still, it may be asked, why should we think of freedom as having intrinsic value in this way? The Oakeshottian answer would be that if we are unsure as to whether we will get to our destination and if we are unsure as to whether we will find the destination meaningful if and when we do arrive there, then we might as well enjoy the journey. One way of making do with a world where there is no guarantee that we will indeed achieve what we want or that what we do achieve will be of lasting significance, is to orient ourselves towards the value of trying to achieve in-of-itself. It is in this regard that cultivating the freedom inherent in agency can help us make do with our lack of collective control over the world. Perhaps looking for meaning in the exercise of agency itself rather than its results or achievements might also render more palatable the absurdity of the human condition. Of course, the cultivation of such an attitude is easier said than done. And perhaps these concluding remarks are suffused with a slightly excessive idealism. Still, given that the argument of this thesis has deflationary implications for certain other ideals, perhaps it is not an undesirable note on which to end.

---

<sup>654</sup> (Oakeshott 2003, 236)

## Bibliography

- Abizadeh, Arash. 2002. 'Does Liberal Democracy Presuppose a Cultural Nation? Four Arguments'. *American Political Science Review* 96(3): 495–509.
- Ackerman, Bruce., and James Fishkin. 2004. *Deliberation Day*. New Haven, Conn.: Yale University Press.
- Acton, John. 1909. 'On Nationality'. In *History of Freedom*, London: Macmillan and Co.
- Adair, Douglass. 1957. "'That Politics May Be Reduced to a Science": David Hume, James Madison, and the Tenth Federalist'. *Huntington Library Quarterly* 20(4): 343–60.
- Alves, André Azevedo, and John Meadowcroft. 2014. 'Hayek's Slippery Slope, the Stability of the Mixed Economy and the Dynamics of Rent Seeking'. *Political Studies* 62(4): 843–61.
- Anderson, Benedict. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London; New York: Verso.
- Arneson, Richard J. 2004. 'Democracy Is Not Inherently Just'. In *Justice and Democracy*, eds. Keith Dowding, Robert E. Goodin, and Carole Pateman. Cambridge: Cambridge University Press, 40–58.
- Arrow, Kenneth J. 1950. 'A Difficulty in the Concept of Social Welfare'. *Journal of Political Economy* 58(4): 328–46.
- Bagg, Samuel. 2018. 'The Power of the Multitude: Answering Epistemic Challenges to Democracy'. *American Political Science Review* 112(4): 891–904.
- Beito, David T. 2000. *From Mutual Aid to the Welfare State: Fraternal Societies and Social Services, 1890-1967*. Chapel Hill: University of North Carolina Press.
- Bevir, Mark, ed. 2013. *Theory of Governance*. Berkeley: Univ. of California Press.
- Bicchieri, Cristina. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Bohman, James. 1997. 'Deliberative Democracy and Effective Social Freedom: Capabilities, Resources and Opportunities'. In *Deliberative Democracy: Essays on Reason and Politics*, eds. James Bohman and William Rehg. Cambridge, Mass: MIT Press, 321–48.
- . 2005. 'From Demos to Demoi: Democracy across Borders'. *Ratio Juris* 18(3): 293–314.
- Bratman, Michael E. 1999. *Faces of Intention: Selected Essays on Intention and*

- Agency*. Cambridge: Cambridge University Press.
- Brennan, Geoffrey, and Loren E. Lomasky. 1997. *Democracy and Decision: The Pure Theory of Electoral Preference*. Cambridge: Cambridge University Press.
- Brennan, Jason. 2007. 'Rawls' Paradox'. *Constitutional Political Economy* 18(4): 287–99.
- . 2014a. 'How Smart Is Democracy? You Can't Answer That Question a Priori'. *Critical Review* 26(1–2): 33–58.
- . 2014b. 'Bleg: Where Are the "Markets Suck" Libertarians?' Bleeding Heart Libertarians. <https://bleedingheartlibertarians.com/2014/07/bleg-where-are-the-markets-suck-libertarians/> (September 8, 2019).
- . 2016. *Against Democracy*. Princeton: Princeton University Press.
- . 2018. 'Private Governance and the Three Biases of Political Philosophy'. *The Review of Austrian Economics* 31(2): 235–43.
- Brennan, Jason, and John Tomasi. 2012. "Classical Liberalism". *The Oxford Handbook of Political Philosophy*, ed. David Estlund. Oxford University Press. <http://oxfordhandbooks.com/view/10.1093/oxfordhb/9780195376692.001.0001/oxfordhb-9780195376692-e-6> (June 3, 2019).
- Buchanan, James, and Gordon Tullock. 1999. *The Calculus of Consent*. Liberty Fund: Indianapolis
- Bull, Hedley. 2012. *The Anarchical Society*. Palgrave Macmillan.
- Caldwell, Bruce. 2008. *Hayek's Challenge: An Intellectual Biography of F.A. Hayek*. Chicago: Univ. of Chicago Press.
- Caplan, Bryan Douglas. 2008. *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton, NJ: Princeton University Press.
- Cavanaugh, William T. 2011. *Migrations of the Holy: God, State, and the Political Meaning of the Church*. Grand Rapids, Mich: William B. Eerdmans Pub. Co.
- Christiano, Thomas. 2008. *The Constitution of Equality*. Oxford University Press. <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198297475.001.0001/acprof-9780198297475> (July 28, 2019).
- Christmas, Billy, and Jason Byas. 2019. 'Methodological Anarchism'. In *The Routledge Handbook of Anarchy and Anarchist Thought*, eds. Gary Chartier and Chad Van Shoelandt. London: Routledge.
- Cohen, G. A. 2003. 'Facts and Principles'. *Philosophy & Public Affairs* 31(3): 211–45.
- . 2009. *Why Not Socialism?* Princeton: Princeton University Press.

- Cohen, Joshua. 1997. 'Deliberation and Democratic Legitimacy'. In *Deliberative Democracy: Essays on Reason and Politics*, eds. James Bohman and William Rehg. Cambridge, Mass: MIT Press, 67–93.
- Constant, Benjamin. 1819. 'The Liberty of Ancients Compared with That of Moderns'. *Liberty Fund: Online Library of Liberty*.  
<https://oll.libertyfund.org/titles/constant-the-liberty-of-ancients-compared-with-that-of-moderns-1819> (May 24, 2019).
- Converse, Philip E. 2006. 'The Nature of Belief Systems in Mass Publics'. *Critical Review* 18(1–3): 1–74.
- Cowen, Nick, and Tranditis, Aris 2019. 'Hayek versus Trump: The Radical Right's Road to Serfdom'. *Polity* (forthcoming).
- Dahl, Robert. 1991. *Democracy and Its Critics*. 12. print. New Haven: Yale Univ. Press.
- .2005. *Who Governs?* New Haven and London: Yale University Press.
- Davies, Stephen. 1997. 'Two Conceptions of Welfare: Voluntarism and Incorporationism'. *Social Philosophy and Policy* 14(2): 39–68.
- DeCanio, Samuel. 2014. 'Democracy, the Market, and the Logic of Social Choice'. *American Journal of Political Science* 58(3): 637–52.
- Dryzek, John S. 2010. *Foundations and Frontiers of Deliberative Governance*. Oxford University Press.  
<http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199562947.001.0001/acprof-9780199562947> (July 23, 2019).
- Easterly, William. 2007. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done so Much Ill and so Little Good*. Oxford; New York: Oxford University Press.
- Ellickson, Robert C. 1991. *Order Without Law: How Neighbors Settle Disputes*. Cambridge, Mass: Harvard University Press.
- Epstein, Richard Allen. 1995. *Simple Rules for a Complex World*. Cambridge, MA: Harvard University Press.
- Fink, Alexander. 2012. 'The Hanseatic League and the Concept of Functional Overlapping Competing Jurisdictions: The Hanseatic League and the Concept of FOCJ'. *Kyklos* 65(2): 194–217.
- Fishkin, James S. 2003. 'Consulting the Public Through Deliberative Polling'. *Journal of Policy Analysis and Management* 22(1): 128–33.
- Foucault, Michel. 1991. 'Governmentality'. In *The Foucault Affect: Studies in Governmentality*, Chicago: The University of Chicago Press.

- Freeman, Samuel Richard. 2007. *Rawls*. London; New York: Routledge.
- Friedman, Jeffrey. 2006. 'Democratic Competence in Normative and Positive Theory: Neglected Implications of "The Nature of Belief Systems in Mass Publics"'. *Critical Review* 18(1–3): 1–43.
- . 2009. 'A Crisis of Politics, Not Economics: Complexity, Ignorance, and Policy Failure'. *Critical Review* 21(2–3): 127–83.
- Gamble, Andrew. 2007. *Hayek: The Iron Cage of Liberty* Cambridge: Polity Press.
- Gaus, Gerald. 2009. 'The Place of Religious Belief in Public Reason Liberalism'. In *Multiculturalism and Moral Conflict*, Routledge innovations in political theory, eds. Maria Dimova-Cookson and Peter M. R. Stirk. Milton Park, Abingdon, Oxon ; New York, NY: Routledge.
- . 2011a. 'A Tale of Two Sets: Public Reason in Equilibrium'. *Public Affairs Quarterly* 25(4): 305–25.
- . 2011b. 'Between Discovery and Choice: The General Will in a Diverse Society'. *Contemporary Readings in Law and Social Justice* 3(2): 70–95.
- . 2012a. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511780844> (April 17, 2019).
- . 2012b. 'Justification, Choice and Promise: Three Devices of the Consent Tradition in a Diverse Society'. *Critical Review of International Social and Political Philosophy* 15(2): 109–27.
- . 2015a. 'The Egalitarian Species'. *Social Philosophy and Policy* 31(2): 1–27.
- . 2015b. 'On Being Inside Social Morality and Seeing It'. *Criminal Law and Philosophy* 9(1): 141–53.
- . 2018. 'The Complexity of a Diverse Moral Order'. *Georgetown Journal of Law & Public Policy* 16(S): 645–80.
- . 2019. *The Tyranny of the Ideal: Justice in a Diverse Society*. Princeton University Press: Princeton <http://dx.doi.org/10.23943/princeton/9780691183428.001.0001> (May 30, 2019).
- Gaus, Gerald, and Shaun Nichols. 2017. 'Moral Learning in the Open Society: The Theory and Practice of Natural Liberty'. *Social Philosophy and Policy* 34(1): 79–101.
- Geuss, Raymond. 2008. *Philosophy and Real Politics*. Princeton: Princeton University Press.



- Gilbert, Margaret. 1992. *On Social Facts*. Princeton, NJ: Princeton Univ. Press.
- Graeber, David. 2015. *The Utopia of Rules: On Technology, Stupidity, and the Secret Joys of Bureaucracy*. Brooklyn; London: Melville House.
- Gray, John. 2000. *Two Faces of Liberalism*. Polity Press.
- Green, David G. 1993. *Reinventing Civil Society: The Rediscovery of Welfare Without Politics*. London: IEA Health and Welfare Unit.
- Guerrero, Alexander A. 2014. 'Against Elections: The Lottocratic Alternative: Against Elections: The Lottocratic Alternative'. *Philosophy & Public Affairs* 42(2): 135–78.
- Gutmann, Amy, and Dennis Thompson. 1998. *Democracy and Disagreement*. Cambridge; London: The Belknap Press of Harvard University Press.
- Habermas, Jürgen. 1983. *Moral Consciousness and Communicative Action*. Cambridge: Polity.
- . 1995. 'Reconciliation Through the Public Use of Reason: Remarks on John Rawls's Political Liberalism'. *The Journal of Philosophy* 92(3): 109.
- . 1996. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge, Mass: MIT Press
- . 2001. 'Constitutional Democracy: A Paradoxical Union of Contradictory Principles?' *Political Theory* 29(6): 766–81.
- Hardin, Russell. 1999. *Liberalism, Constitutionalism, and Democracy*. Oxford University Press.  
<http://www.oxfordscholarship.com/view/10.1093/0198290845.001.0001/acprof-9780198290841> (April 1, 2019).
- Hayek, Friedrich. 1945. 'The Use of Knowledge in Society'. *The American Economic Review* 35(4): 519–30.
- . 1948. 'The Economic Conditions of Interstate Federalism'. In *Individualism and the Economic Order*, The University of Chicago Press, 255–72.
- . 1949. 'The Intellectuals and Socialism'. *The University of Chicago Law Review* 16(3):
- . 1955. 'Degrees of Explanation'. *The British Journal for the Philosophy of Science* 6(23): 209–25.
- . 1967a. 'Rules, Perception and Intelligibility'. In *Studies in Philosophy, Politics and Economics*, Chicago: University of Chicago Press, 43–65.
- . 1967b. 'The Theory of Complex Phenomena'. In *Studies in Philosophy,*

- Politics and Economics*, Chicago: University of Chicago Press, 22–42.
- . 1967c. ‘The Road to Serfdom after Twelve Years\*’. In *Studies in Philosophy, Politics And Economics*, Touchstone, 216–28.
- . 1982a. *Law, Legislation and Liberty: Vol. 1 Rules and Order*. 3rd ed. Abingdon: Routledge.
- . 1982b. *Law, Legislation and Liberty: Vol. 2 The Mirage of Social Justice*. 3rd ed. Abingdon: Routledge.
- . 1982c. *Law, Legislation and Liberty: Vol. 3 The Political Order of a Free People*. 3rd ed. Abingdon: Routledge. 417–33.
- . 1987 [1952]. *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*. Reprint. London: Routledge & Kegan Paul.
- . 1988. *The Fatal Conceit: The Errors of Socialism*. London: Routledge.
- . 2001 [1944]. *The Road to Serfdom*. London: Routledge.
- . 2006 [1960]. *The Constitution of Liberty*. London: Routledge.
- Hegel, Georg Wilhelm Friedrich, 2011 [1896]. *Elements of the Philosophy of Right*. Cambridge: Cambridge Univ. Press.
- Held, David. 2007. *Democracy and the Global Order: From the Modern State to Cosmopolitan Governance*. Cambridge: Polity.
- Hendriks, Carolyn M. 2006. ‘Integrated Deliberation: Reconciling Civil Society’s Dual Role in Deliberative Democracy’. *Political Studies* 54(3): 486–508.
- Horton, John. 2003. ‘Rawls, Public Reason and the Limits of Liberal Justification’. *Contemporary Political Theory* 2(1): 5–23.
- Huemer, Michael. 2013. *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey*. Hampshire; New York: Palgrave Macmillan.
- Humboldt, Wilhelm von. 1854. *The Sphere and Duties of Government (The Limits of State Action)*. London: John Chapman and King Willian Street.  
<https://oll.libertyfund.org/titles/humboldt-the-sphere-and-duties-of-government-1792-1854> (July 16, 2019).
- Hume, David. 1777a. ‘Of Parties in General’. In *Essays Moral, Political, Literary*, Indianapolis: Liberty Fund. <https://oll.libertyfund.org/titles/hume-essays-moral-political-literary-lf-ed> (March 27, 2019).
- . 1777b. ‘Of The Parties of Great Britain’. In *Essays Moral, Political, Literary*, Indianapolis: Liberty Fund. <https://oll.libertyfund.org/titles/hume-essays->

moral

- . 1777c. 'Idea of a Perfect Commonwealth'. In *Essays Moral, Political, Literary*, Indianapolis: Liberty Fund. <https://oll.libertyfund.org/titles/hume-essays-moral-political-literary-lf-ed> (March 27, 2019).
- . 1777d. 'Of The Original Contract'. In *Essays Moral, Political, Literary*, Indianapolis: Liberty Fund. <https://oll.libertyfund.org/titles/hume-essays-moral-political-literary-lf-ed> (June 2, 2019).
- Ikeda, Sanford. 2003. 'How Compatible Are Public Choice and Austrian Political Economy?' *The Review of Austrian Economics* 16(1): 63–75.
- Johnston, David. 1997. 'Is the Idea of Social Justice Meaningful?' *Critical Review* 11(4): 607–14.
- Kelly, Paul. 2005. *Liberalism*. Cambridge: Polity.
- Kogelmann, Brian. 2019. 'Public Reason's Chaos Theorem'. *Episteme* 16(2): 200–219.
- Kogelmann, Brian, and Stephen G. W. Stich. 2016. 'When Public Reason Fails Us: Convergence Discourse as Blood Oath'. *American Political Science Review* 110(4): 717–30.
- Kothari, S.P., and Rebecca Lester. 2011. 'The Role of Accounting in the Financial Crisis: Lessons for the Future'. SSRN Electronic Journal. <http://www.ssrn.com/abstract=1972354> (September 19, 2019).
- Kukathas, Chandran. 1989. *Hayek and Modern Liberalism*. Oxford [England]; New York: Clarendon Press; Oxford University Press.
- . 2003. *The Liberal Archipelago: A Theory of Diversity and Freedom*. Oxford: Oxford University Press.
- . 2006. 'Hayek and Liberalism'. In *The Cambridge Companion to Hayek*, ed. Edward Feser. New York: Cambridge University Press, 182–208.
- Kurzman, Charles. 2005. *The Unthinkable Revolution in Iran*. Cambridge, Mass. London: Harvard University Press.
- Kymlicka, Will. 1996. *Multicultural Citizenship*. Oxford University Press. <http://www.oxfordscholarship.com/view/10.1093/0198290918.001.0001/acprof-9780198290919> (August 9, 2019).
- Landemore, Hélène. 2017. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton: Princeton University Press
- Larmore, Charles. 1999. 'The Idea of a Life Plan'. *Social Philosophy and Policy* 16(1): 96–112.

- Leeson, Peter. 2008. 'How Important Is State Enforcement for Trade?' *American Law and Economics Review* 10(1): 61–89.
- Levy, Jacob. 2016a. *Rationalism, Pluralism, and Freedom*. Oxford: OUP.
- . 2016b. 'There Is No Such Thing as Ideal Theory'. *Social Philosophy and Policy* 33(1–2): 312–33.
- . 2017. 'Against Fraternity'. In *The Strains of Commitment*, eds. Keith Banting and Will Kymlicka. Oxford University Press, 107–24.  
<http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198795452.001.0001/acprof-9780198795452-chapter-4> (September 2, 2019).
- . 2018. 'Political Libertarianism'. In *The Cambridge Handbook of Classical Liberal Thought*, ed. Todd Henderson. Cambridge: Cambridge University Press.
- Lewis, Paul. 2012. 'Emergent Properties in the Work of Friedrich Hayek'. *Journal of Economic Behavior & Organization* 82(2–3): 368–78.
- . 2016. 'The Emergence of "Emergence" in the Work of F. A. Hayek: A Historical Analysis'. *History of Political Economy* 48(1): 111–50.
- Lindsey, Brink. 2006. 'Liberalarians'. *The New Republic*.  
<https://newrepublic.com/article/64443/liberalarians> (August 5, 2019).
- Lindblom, Charles E. 1959. 'The Science of "Muddling Through"'. *Public Administration Review* 19(2): 79.
- Lipsky, Michael. 2010. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*. Russell Sage Foundation.
- List, Christian. 2011. 'The Logical Space of Democracy'. *Philosophy & Public Affairs* 39(3): 262–97.
- List, Christian, and Philip Pettit. 2006. 'Group Agency and Supervenience'. *The Southern Journal of Philosophy* 44(S1): 85–105.
- Lister, Andrew. 2013. 'The "Mirage" of Social Justice: Hayek Against (and For) Rawls'. *Critical Review* 25(3–4): 409–44.
- Lomasky, Loren E. 2011. 'Liberty After Lehman Brothers'. *Social Philosophy and Policy* 28(02): 135–65.
- Lukes, Steven. 1997. 'Social Justice: The Hayekian Challenge'. *Critical Review* 11(1): 65–80.
- Macedo, Stephen. 2004. 'What Self-Governing Peoples Owe to One Another: Universalism, Diversity, and the Law of Peoples'. *Fordham Law Review* 72: 19.

- Machiavelli, Niccolò. 2004. *The Prince*. London: Penguin Books.
- MacIntyre, Alasdair C. 2007. *After Virtue: A Study in Moral Theory*. Notre Dame, Ind: University of Notre Dame Press.
- Mack, Eric. 2013. 2018. *Libertarianism*. Oxford: Oxford University Press.
- Madison, James. 2003a. 'The Federalist No.10'. In *Hamilton, Madison and Jay: The Federalist with Letters of 'Brutus'*, Cambridge: Cambridge University press, 40–46.
- . 2003b. 'The Federalist No.51'. In *Hamilton, Madison and Jay: The Federalist with Letters of 'Brutus'*, Cambridge: Cambridge University press, 251–55.
- Mansbridge et. al., Jane. 2012. 'A Systemic Approach to Deliberative Democracy'. In *Deliberative Systems: Deliberative Democracy at the Large Scale*, Theories of institutional design, eds. John Parkinson and Jane J. Mansbridge. Cambridge: Cambridge University Press, 1–27.
- Margalit, Avishai, and Joseph Raz. 1990. 'National Self-Determination'. *Journal of Philosophy* 87(9): 439–61.
- Mauss, Marcel. 2002. *The Gift: The Form and Reason for Exchange in Archaic Societies*. London and New York: Routledge.
- McCabe, David. 2010. *Modus Vivendi Liberalism*. Cambridge University Press.  
/core/books/modus-vivendi-liberalism/CDBF4F7F169F63E6949357B155DA111F (March 29, 2019).
- McCloskey, Deirdre N. 2011. 'Hobbes, Rawls, Nussbaum, Buchanan, and All Seven of the Virtues'. *Journal des Economistes et des Etudes Humaines* 17(1).  
<https://www.degruyter.com/view/j/jeeh.2011.17.issue-1/jeeh.2011.17.1.1249/jeeh.2011.17.1.1249.xml> (July 17, 2019).
- Mill, John Stuart. 1861. *Considerations on Representative Government*. Dodo Press.
- Miller, David. 1997. *On Nationality*. Oxford University Press.  
<http://www.oxfordscholarship.com/view/10.1093/0198293569.001.0001/acprof-9780198293569> (April 30, 2019).
- . 2003. *Principles of Social Justice*. Cambridge, Mass.: Harvard Univ. Press.
- . 2007. *National Responsibility and Global Justice*. Oxford University Press.  
<http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199235056.001.0001/acprof-9780199235056> (April 30, 2019).
- Mises, Ludwig Von. 1920. *Economic Calculation in the Socialist Commonwealth*. Ludwig Von Mises Institute.  
[https://mises.org/sites/default/files/Economic%20Calculation%20in%20the%20Socialist%20Commonwealth\\_Vol\\_2\\_3.pdf](https://mises.org/sites/default/files/Economic%20Calculation%20in%20the%20Socialist%20Commonwealth_Vol_2_3.pdf) (September 3, 2019).

- Moore, Margaret. 2015. *A Political Theory of Territory*. Oxford University Press.  
<http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780190222246.001.0001/acprof-9780190222246> (April 30, 2019).
- Müller, Jan-Werner. 2015. 'What, If Anything, Is Wrong with Hayek's Model Constitution?' In *Law, Liberty and State: Oakeshott, Hayek and Schmitt on the Rule of Law*, ed. David Dyzenhaus. Cambridge University Press.  
[/core/books/law-liberty-and-state/what-if-anything-is-wrong-with-hayeks-model-constitution/68418CF6614EF1B8ADE6DBDCCC0343F0/core-reader](http://www.oxfordscholarship.com/core/books/law-liberty-and-state/what-if-anything-is-wrong-with-hayeks-model-constitution/68418CF6614EF1B8ADE6DBDCCC0343F0/core-reader)  
 (April 8, 2019).
- Murphy, Liam, and Thomas Nagel. 2002. *The Myth of Ownership*. Oxford University Press.  
<http://www.oxfordscholarship.com/view/10.1093/0195150163.001.0001/acprof-9780195150162> (June 7, 2019).
- Nagel, Thomas. 1971. 'The Absurd'. *The Journal of Philosophy* 68(20): 716.
- . 2005. 'The Problem of Global Justice'. *Philosophy & Public Affairs* 33(2): 113–47.
- Nisbet, Robert A. 1953. *The Quest for Community: A Study in the Ethics of Order and Freedom*. New critical ed. Wilmington, Del: ISI Books.
- North, Douglass C., John Joseph Wallis, and Barry R. Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. <https://www.cambridge.org/core/books/violence-and-social-orders/F0EA15A67E790214408A7485DBC70F0D> (March 28, 2019).
- North, Douglass C., and Barry R. Weingast. 1989. 'Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England'. *The Journal of Economic History* 49(4): 803–32.
- Nozick, Robert. 2012 [1974]. *Anarchy, State, and Utopia*. Nachdr. Malden, MA: Blackwell.
- Oakeshott, Michael. 1991. *Rationalism in Politics and Other Essays*. New and expanded ed. Indianapolis: Liberty Fund.
- . 2003. *On Human Conduct*. Oxford: Oxford University Press.
- Olson, Mancur. 2003 [1971]. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, Mass.: Harvard Univ. Press.
- O'Neill, Onora. 1996. *Towards Justice and Virtue: A Constructive Account of Practical Reasoning*. Cambridge: Cambridge University Press.  
<http://ebooks.cambridge.org/ref/id/CBO9780511621239> (July 29, 2019).
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge ; New York: Cambridge University Press.

- Ostrom, Vincent. 1972. 'Polycentricity'. In Washington Hilton Hotel, Washington, D.C. <http://dlc.dlib.indiana.edu/dlc/handle/10535/3763> (July 3, 2019).
- Pennington, Mark. 2003. 'Hayekian Political Economy and the Limits of Deliberative Democracy'. *Political Studies* 51(4): 722–39.
- . 2011. *Robust Political Economy: Classical Liberalism and the Future of Public Policy*. Cheltenham, UK; Northampton, MA, USA: Edward Elgar.
- . 2015. 'Constitutional Political Economy and Austrian Economics'. In *The Oxford Handbook of Austrian Economics*, eds. Christopher J. Coyne, Peter J. Boettke, Oxford University Press. <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199811762.001.0001/oxfordhb-9780199811762-e-20> (September 4, 2019).
- Pettit, Philip. 1997. *Republicanism: A Theory of Freedom and Government*. Oxford: New York: Clarendon Press; Oxford University Press.
- Pettit, Philip, and David Schweikard. 2006. 'Joint Actions and Group Agents'. *Philosophy of the Social Sciences* 36(1): 18–39.
- Pevnick, Ryan. 2014. *Immigration and the Constraints of Justice: Between Open Borders and Absolute Sovereignty*. Cambridge: Cambridge University Press.
- Plant, Raymond. 1994. 'Hayek on Social Justice: A Critique'. In *Hayek, Co-Ordination and Evolution*, eds. Rudy van Zijp and Jack Birner. Routledge. <http://www.crcnetbase.com/doi/10.4324/9780203004241.ch8> (April 12, 2019).
- Polanyi, Michael. 2005. *Personal Knowledge Towards a Post-Critical Philosophy*. London: Routledge.
- Putnam, Robert D. 2001. *Bowling Alone: The Collapse and Revival of American Community*. New York, NY: Simon & Schuster.
- Quong, Jonathan. 2011. *Liberalism without Perfection*. Oxford: Oxford University Press.
- Rawlins, Michael D, and Anthony J Culyer. 2004. 'National Institute for Clinical Excellence and Its Value Judgments'. *British Medical Journal* 329(7459): 224–27.
- Rawls, John. 1963. 'The Sense of Justice'. *The Philosophical Review* 72(3): 281.
- . 1974. 'The Independence of Moral Theory'. *Proceedings and Addresses of the American Philosophical Association* 48: 5–22.
- . 1980. 'Kantian Constructivism in Moral Theory'. *The Journal of Philosophy* 77(9): 515.

- . 1996. *Political Liberalism*. 2nd ed. New York: Columbia University Press.
- . 1997. 'The Idea of Public Reason Revisited'. *The University of Chicago Law Review* 64(3): 44.
- . 1999. *A Theory of Justice: Revised Edition*. 2nd ed. Cambridge: The Belnap Press of Harvard University Press.
- . 2001. *Justice as Fairness: A Restatement*. Cambridge and London: Harvard University Press.
- . 2003. *The Law of Peoples*. Cambridge, Mass.: Harvard Univ. Press.
- Riker, William H. 1988. *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice*. Prospect Heights, Ill: Waveland Press.
- Roberts, Paul Craig. 2002. 'My Time with Soviet Economics'. *The Independent Review* 7(2): 259–64.
- Roberts, Russell D. 2019. *Gambling with Other People's Money: How Perverse Incentives Caused the Financial Crisis*. Stanford, California: Hoover Institution Press, Stanford University.
- Rousseau, Jean-Jacques. 1968 [1762]. *The Social Contract*. ed. Maurice William Cranston. Harmondsworth [Middlesex: Penguin).
- Ryle, Gilbert. 1945. 'Knowing How and Knowing That: The Presidential Address'. *Proceedings of the Aristotelian Society* 46: 1–16.
- Sandel, Michael J. 1998. *Democracy's Discontent: America in Search of a Public Philosophy*. Cambridge, Mass.: Belknap Press of Harvard Univ. Press.
- Salter, Alexander William, and Andrew T. Young. 2019. 'Polycentric Sovereignty: The Medieval Constitution, Governance Quality, and the Wealth of Nations'. *Social Science Quarterly* 100(4): 1241–53.
- Schmidtz, David. 2006. *Elements of Justice*. Cambridge; New York: Cambridge University Press.
- . 2010. 'Property and Justice'. *Social Philosophy and Policy* 27(1): 79–100.
- . 2016. 'Friedrich Hayek'. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/win2017/entries/friedrich-hayek/> (July 31, 2019).
- Schmitt, Carl. 2005. *Political Theology: Four Chapters on the Concept of Sovereignty*. University of Chicago Press ed. Chicago: University of Chicago Press.
- Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the*



- Human Condition Have Failed*. New Haven: Yale University Press.
- . 2012. *Two Cheers for Anarchism: Six Easy Pieces on Autonomy, Dignity, and Meaningful Work and Play*. Princeton: Princeton University Press
- Searle, John. 1990. 'Collective Intentions and Actions'. In *Intentions in Communication*, System Development Foundation Benchmark series, ed. Philip R. Cohen. Cambridge, Mass.: MIT Press, 401–15.
- Sen, Amartya. 1981. *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford : New York: Clarendon Press; Oxford University Press.
- . 2009. *The Idea of Justice*. London: Allen Lane.
- Shearmur, Jeremy. 1996. *Hayek and After: Hayekian Liberalism as a Research Programme*. London; New York: Routledge.
- Simmons, A. John. 1976. 'Tacit Consent and Political Obligation'. *Philosophy & Public Affairs* 5(3): 274–91.
- . 1999. 'Justification and Legitimacy'. *Ethics* 109(4): 739–71.
- Skarbek, David. 2011. 'Governance and Prison Gangs'. *American Political Science Review* 105(4): 702–16.
- Somin, Ilya. 2013. *Democracy and Political Ignorance: Why Smaller Government Is Smarter*. Stanford, California: Stanford Law Books, an imprint of Stanford University Press.
- Spruyt, Hendrik. 1994. *The Sovereign State and Its Competitors: An Analysis of Systems Change*. Princeton, N.J: Princeton University Press.
- Stilz, Anna. 2009. *Liberal Loyalty: Freedom, Obligation, and the State*. Princeton: Princeton University Press.
- . 2015. 'Decolonization and Self-Determination'. *Social Philosophy and Policy* 32(1): 1–24.
- . 2016. 'The Value of Self-Determination'. In *Oxford Studies in Political Philosophy*, OUP.
- Stringham, Edward Peter. 2015. *Private Governance: Creating Order in Economic and Social Life*. Oxford University Press.  
<http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199365166.001.0001/acprof-9780199365166> (June 4, 2019).
- Tamir, Yael. 1993. *Liberal Nationalism*. Princeton, N.J: Princeton University Press.
- Taylor, Dianna. 2009. 'Normativity and Normalization'. *Foucault Studies* (7): 45.

- Taylor, John B. 2009. 'Economic Policy And The Financial Crisis: An Empirical Analysis Of What Went Wrong'. *Critical Review* 21(2–3): 341–64.
- Tebble, Adam. 2003. 'Does Inclusion Require Democracy?' *Political Studies* 51: 197–214.
- . 2016a. *Epistemic Liberalism: A Defence*. Abingdon: Routledge.
- . 2016b. 'On the Circumstances of Justice'. *European Journal of Political Theory* 0(0): 1–23.
- Thrasher, John, and Kevin Vallier. 2013. 'The Fragility of Consensus: Public Reason, Diversity and Stability'. *European Journal of Philosophy* 23(4): 933–54.
- Tilly, Charles. 1992. *Coercion, Capital, and European States, AD 990-1992*. Cambridge, MA: Blackwell.
- Tocqueville, Alexis de. 2003 [1840]. *Democracy in America: And Two Essays on America*. London: Penguin.
- Tomasi, John. 2012. 'Democratic Legitimacy and Economic Liberty'. *Social Philosophy and Policy* 29(1): 50–80.
- . 2013. *Free Market Fairness*. Princeton and Oxford: Princeton University Press.
- Tuomela, Raimo. 2005. 'We-Intentions Revisited'. *Philosophical Studies* 125(3): 327–69.
- UK Government website. 2019. 'Departments, Agencies and Public Bodies'. <https://www.gov.uk/government/organisations> (April 15, 2019).
- Valentini, Laura. 2012. 'Ideal vs. Non-Ideal Theory: A Conceptual Map: Ideal vs Non-Ideal Theory'. *Philosophy Compass* 7(9): 654–64.
- . 2013. 'Justice, Disagreement and Democracy'. *British Journal of Political Science* 43(1): 177–99.
- Van Creveld, Martin. 1999. *The Rise and Decline of the State*. Cambridge [England]; New York: Cambridge University Press.
- Vaughn, Karen I. 1999. 'Hayek's Theory of the Market Order as an Instance of the Theory of Complex, Adaptive Systems'. *Journal des Économistes et des Études Humaines* 9(2–3). <https://www.degruyter.com/view/j/jeeh.1999.9.issue-2-3/jeeh-1999-2-304/jeeh-1999-2-304.xml> (April 10, 2019).
- Velleman, J. David. 1997. 'How To Share An Intention'. *Philosophy and Phenomenological Research* 57(1): 29.

- Waldron, Jeremy. 2004. *Law and Disagreement*. Oxford: Clarendon Press.
- Weber, Eugen. 2007. *Peasants into Frenchmen: The Modernization of Rural France ; 1870 - 1914*. Nachdr. Stanford, Calif: Stanford Univ. Press.
- Weingast, Douglas. 1995. 'The Economic Role of Political Institutions: Market-Preserving Federalism and Economic Development'. *The Journal of Law, Economics, and Organization* 11(1): 1–31.
- Weithman, Paul. 2011. *Why Political Liberalism?: On John Rawls's Political Turn*. Oxford University Press.  
<http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195393033.001.0001/acprof-9780195393033> (April 1, 2019).
- Wendt, Alexander. 1992. 'Anarchy Is What States Make of It: The Social Construction of Power Politics'. *International Organization* 46(2): 391–425.
- Wilkinson, Will. 2006. 'Is Rawlsianism the Future?' *Cato Institute*.  
<https://www.cato.org/blog/rawlsianism-future> (August 5, 2019).
- Williams, Bernard. 2008. *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Princeton, N.J.: Princeton Univ. Press.
- Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. 2nd ed. Oxford: Basil Blackwell.
- Young, Iris Marion. 1990. *Justice and the Politics of Difference*. Princeton and New Jersey: Princeton University Press.
- Zhu, Jiangnan, and Dong Zhang. 2017. 'Weapons of the Powerful: Authoritarian Elite Competition and Politicized Anticorruption in China'. *Comparative Political Studies* 50(9): 1186–1220.