

Giulio Fornaroli

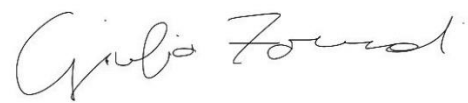
**Responding to Disagreement in Liberal
Societies: Legitimacy, Respect, Toleration**

University College London – School of Public Policy (Department of Political Science)

Thesis for a PhD degree in political science

I, Giulio Fornaroli, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

London, 11th September 2018.

A handwritten signature in black ink that reads "Giulio Fornaroli". The signature is written in a cursive style with a large initial 'G' and 'F'.

Abstract

Consider the nature of disagreement about the legal permissibility of abortion within a democratic society. For the purposes of this thesis, it has two important features: it is a disagreement about matters of value, and it takes place among people who, because it is a democratic state, are all entitled to have a say (in some form) on the matter. This means some participants in the debate will have normative requirements imposed on them to which they have reasons to object. What is the correct moral response to this fact as a matter of political philosophy?

One of the most influential currents in contemporary political philosophy, political liberalism, holds that disagreements of this kind can be addressed successfully within a single polity if the state is legitimate, i.e., if it is capable of providing a set of *public* reasons justifying its authority. Because the reasons are public and accessible to people holding a plurality of different but reasonable conceptions of value, reasonable citizens holding these conceptions will face normative requirements that they will believe are acceptable. Chapters 1-3 of this thesis show this optimistic expectation about what citizen's reasonableness entails for legitimacy and disagreement are ill-founded.

I restore some optimism in the second half of the dissertation by arguing that there are key values we can appeal to in plausibly responding to the existence of value-disagreement. These include the central liberal notions of *equal respect* and *toleration*, theories of which I set out in Chapters 4 and 5 respectively. I elaborate a conception of each concept based on the insight that they are fundamental attitudes citizens need to develop in order for a just society to exist and remain stable. I conclude, in Chapter 6, by showing that this insight can help us design good policies in responding to the fact of value-disagreement when dealing with how we should accommodate conscientious reasons for rejecting a state requirement.

Impact Statement

The current political debate on such questions as immigration, conscientious objections, or minority rights, often appears as a clash of completely incompatible views. On the one hand, you find those that affirm that diversity is a value, and that the inclusion of more groups and ideas within a single society, or the new opportunities that are offered to groups that were previously marginalised, can only help the society flourish, at least in the long term. On the other, there are those that see diversity as the most significant threat on the preservation of the political order, and believe that, again in the long term, a society that does not pose strict limits to toleration is going to fall victim to its own relaxed attitude.

This is of course a simplification, but constitutes an interesting starting point for this work. In this thesis, I do not embrace any of the two incompatible views, but I use the conceptual instruments of political theory to demonstrate something that advocates of both positions might find interesting.

Some prominent political theorists before me have attempted to design a solution, but, I try to show here, they have not been successful. I focus in particular on political liberals, who claim that no matter how diverse a society here, if the views people embrace pass a threshold of 'reasonableness', then living together is not that complicated because we will easily converge or find a consensus on a minimal conception of justice that can help us regulate our society. We will keep disagreeing about what ultimately matters in human life, but we will have procedures to settle our disagreement in the political arena.

This solution is reasonable, but, I argue, hopeless. There are simply too many problems that this model gives rise to. For example, how do we determine a threshold of reasonableness? Or, what does it mean, in practice, to converge on or find a

consensus on certain minimal principles of justice? It seems that a solution that requires further problems to be addressed is not, after all, a great solution.

I suggest something different. Instead of focusing on the possibility of finding a consensus on principles of justice, or on justifying each other the norms we use to regulate society, we should instead follow certain virtuous attitudes when dealing with diversity in a democratic society. The first attitude is equal respect: we should respect others' willingness and capacity to form their own conception of what matters in their life. We should not try to hinder them in this fundamental pursuit, for example by misconstruing or manipulating what they are saying when they advance their claims publicly. The second attitude is toleration: we should accept that other we disagree with are as entitled as we are to try to define what our society will eventually look like. That means that we should also accept, and not too grudgingly, that our ideal model of the perfect society is not going to be realised.

All of this is good, someone might counter, but aren't you taking an overly idealised view of politics? I reply to this in the very last parts of the work. I show there that a just society can do a lot to promote these virtuous attitudes, at least in the sense that it can render them convenient options. Hence, I conclude, the concrete justness of the public institutions we live in, and not their abstract legitimacy, is what we should focus on, when we deal with diversity and disagreement.

Table of Contents

Introduction	9
0.1 Introducing Disagreement among Moral Peers – Summary of the Work.....	9
0.1.1 The Point of the Work.....	9
0.1.2 Summary of the Work.....	12
0.2 Why Disagreement? Which Disagreement?	18
0.2.1 The Question of Public Justification in Political and Justificatory Liberalism	18
0.2.2 Respect, Vulnerability and Stability: Why Liberal Theories of Justice Cannot Ignore Disagreement.....	25
0.3 Limits and Aspirations	30
0.3.1 In Defence of (Ideal) Theory.....	30
1. Political Liberalism and the Question of Political Legitimacy	40
1.1 A Model of Political Legitimacy.....	42
1.1.1 Rejecting the ‘Anarchical Challenge’ Model	42
1.1.2 Two Senses of Political Legitimacy.....	47
1.2 Legitimacy as Territorial Sovereignty and Disagreement.....	59
1.2.1 On an Account of Legitimacy as Territorial Sovereignty	59
1.2.2 On a Standard of Territorial Sovereignty.....	67
1.2.3 Conventions and Reasonable Disagreement	76
2. The Failure of Public Justification	81
2.1 Two Models of Public Justification.....	84
2.2 A Critique of Convergence	92
2.2.1 On the Epistemological Demandingness of Convergence	92
2.2.2 Convergence and the Fetishism of Public Justification	96
2.3 Reasonable Consensus and the Problem of Interpretation	100
2.3.1 Preliminary Observations about Reasonableness.....	100
2.3.2 Reasonable Consensus and the Problem of Interpretive Disagreement.....	107

3. Reasonableness and the Problematic Epistemology of Political Liberalism	119
3.1 A Critique of the Appeal to Reasonableness in Public Justification	122
3.1.1 Introducing the Problem.....	122
3.1.2. Reasonableness and the Problem of Scepticism – Rejection of the Epistemic Systems Paradigm.....	126
3.1.3 Reasonableness and the Problem of Scepticism – Rejection of Extreme Permissivism in the Face of Reasonable Disagreement.....	130
3.2 Conclusions on Legitimacy and Disagreement.....	138
4. Equal Respect for Differences: An Oxymoron?	142
4.1 From Kant’s Account to a Kantian Ideal.....	145
4.1.1 Kant’s Original Account and Its Difficulties.....	145
4.1.2 The Problem of the Identification of the Properties on Which Respect Might Supervene.....	148
4.1.3 A Shift in Paradigm: Equal Respect without Empirical (or Metaphysical) Foundations.....	153
4.2 Respect in Circumstances of Disagreement: Substantive Accounts	159
4.2.1 ‘Justificatory Liberalism’ and the Principle of Sincerity.....	159
4.2.2 Respect as ‘Hands-on Recognition’	164
4.2.3 The Problem of Evaluations and the Failure of ‘Hands-on Recognition’	169
4.3 Scepticism, ‘Mediated Respect’ and ‘As-If’ Rationality.....	176
5. Stability as Toleration.....	185
5.1 Public Justification and Stability – A Sceptical Perspective	189
5.1.1 Stability as a Problem of Assurance	189
5.1.2 Scepticism over Public Justification and Assurance	195
5.2 On Political Realism and Its Insight	205
5.3 Stability-Enhancing Toleration.....	215
5.3.1 From Rawls’s ‘Aristotelian Principle’ to Toleration as Equal Entitlement.....	215
5.3.2 The Rejection of Modus Vivendi	219
5.3.3 Toleration as a Virtue.....	224
5.4 Final Reflections on Justice and Stability.....	231

6. The Possibility of Accommodation and the Rejection of the Integrity Paradigm	235
6.1 Liberalism and the Hopeless Valuation of Integrity.....	238
6.2 The Conditional Value of Integrity and Its Failure	248
6.3 The Elusive Role of Integrity.....	255
6.4 Closing Remarks on the Question of Accommodations	265
Closing Remarks	268
7.1 Liberal Pluralism beyond Legitimacy	269
Bibliography	284

Introduction

0.1 Introducing Disagreement among Moral Peers – Summary of the Work

0.1.1 *The Point of the Work*

Imagine the fellow members of an organisation – say, a university department – are stuck in a disagreement regarding a controversial topic. Some members – call them the ‘reformers’ – contest the exposure, in the department’s most prestigious public space, of the pictures of all the department’s previous directors – not unusually, a monochrome group of middle-aged white men, a couple of which said or did things in their life that the contemporary moral gaze would tend to condemn.¹ These ‘reformers’ see the pictures as a representation of academia that sits in sharp contrast to the diversifying spirit that ought to inform academic practice today; they also fear that students coming from under-represented groups (women and people of colour) might find that image discomfoting. Other department members – call them the ‘conservatives’ – do not see the pictures as a celebration of the whiteness and masculinity that used to monopolise academia, nor of every single detail regarding the private and public life of the people portrayed. They see the pictures instead as a symbolic representation of the history of a practice they value in its integrity, despite its containing, like virtually all other practices that have had a significant duration in time, bad examples together with the good.

Political theory can be interested in a scenario of this kind for various reasons. In the present work, however, I intend to highlight just two of them. The first is that the various parties, ‘conservatives’ and ‘reformers’ alike, all feel entitled to have their voice heard and to have an impact in the decision that will be eventually taken.

¹ Let’s say, for example, that one admitted in his memoirs some forms of sexual misconduct towards female students, and another was notoriously prone to casually racist jokes. The example is not entirely fictitious. A similar discussion has recently taken place at UCL Department of Philosophy.

They demand to be treated, and they assume they will be treated, as *peers*. Presumably, one of the justifications for treating all participants to the debates as peers is that they are all (roughly) equally affected by the final decision. But it seems there is a further, and more fundamental, assumption, which is that all participants are, despite their differences, equally capable of deciding autonomously what is the best option, at least from their own perspective. Let's call this form of parity, in need of a better name, 'moral parity'.

The second interesting feature of the scenario is that it involves a disagreement that is centrally about matters of subjective value. We can define subjective value, in a trivial but overarching way, as whatever people take to be good or relevant, beyond the mere strive for survival, in their life; whatever, in slightly different terms, they believe could make their life worth living.

We can then redefine the scenario above as one where a disagreement about matters of (individual or subjective) value takes place among moral peers. What could political theory have to say about this?

An answer comes from political liberalism. Political liberals are concerned with the types of value-disagreement that threaten the stability of a just, democratic society, or, in the words of John Rawls, of a 'well-ordered democratic society'. They believe that the state, when legitimate, can justify its authority in terms that all citizens who hold reasonable (subjective or individual)² sets of values (or, in their jargon, 'reasonable conceptions of the good') see as acceptable. That way, even if citizens keep disagreeing about which values ought to prevail publicly, they will have institutions and procedures, which they all accept, to settle their disputes.

² In the rest of the work, I am going to talk about value-disagreement to always refer to disagreement concerning subjective values.

The first half of the present work will demonstrate that this optimistic picture is, unfortunately, ill-founded. I will gradually attack some central components of the grand project of political liberalism but focus in particular on the belief that a scheme of public justification that is accessible to reasonable citizens is all we need to claim that the state is both stable and legitimate. Depending on how we decide to develop it, public justification is, I will argue, either too demanding a requirement for states in the real world or an inadequate response to the question of value-disagreement. The political liberal focus on political legitimacy as a plausible way to respond to disagreement in politics seems therefore misplaced.

The second half of the work will serve to redress these pessimistic conclusions by showing that justice can still be reconciled with the existence of value-disagreement if political societies are regulated following two fundamental principles that have always been part of the liberal repertoire: equal respect and toleration. Equal respect demands that we recognise in others the same capacity to form, revise and live after an autonomously chosen conception of the good that we would recognise, and value, in ourselves. The principle of equal respect will generate particular moral duties. Those mandate that, in circumstances of value-disagreement and deliberation (as in the scenario above), people adopt a particular *attitude*, namely the willingness to acquire a full and unbiased comprehension of the others' viewpoint.

Toleration corresponds instead to an ascription to fellow citizens of a particular moral status. Precisely, toleration corresponds to a recognition that one's fellows are as entitled as one is to try to shape society according to what they believe are reasonable criteria of justice. Toleration, I will argue, plays the most significant role in the solution to another challenge that the existence of disagreement is generally thought to bring about, the problem of political stability.

I am going to interpret the problem of stability – in line with current scholarship – as consisting in a lack of assurance that citizens will comply with norms that they perceive are in contrast with their interests or values. I will then show that the tolerant attitude will help to provide exactly that type of assurance. Citizens who are happy to consider their fellows entitled to contribute to the definition of their own society do that fully aware of the consequences of this line of conduct. They know that the society might turn out to be slightly different from the ideal model they would love to see realised, and they will not therefore stop complying when their public demands have been slightly frustrated.

The identification of equal respect and toleration as public virtues that citizens ought to develop in a context of endemic and irreducible value-disagreement will lead to the central normative conclusion of the present work, which is that norms and institutions should be designed so that toleration and equal respect become, for as many citizens as possible, convenient and valuable options.

0.1.2 Summary of the Work

This is how the work will develop. I will start, in this very Introduction (§ 0.2), by presenting both the problem of disagreement in the context of political theory and the literature I will mostly engage with in rest of the work. At the end of the section, I will have shown that, from a broadly liberal perspective, we ought to recognise two kinds of reasons to take into account the existence of disagreement. The first set of considerations are deontological and relate to the importance that certain disagreements have for those who take part in them in terms of conscientious involvement. This consideration would be normatively vacuous if it were not coupled with the moral requirement that a political institution owes some form of equal respect to its citizens' conscience.

The second set of considerations is teleological and relates instead to the role disagreement might play in a society's political stability. I will here start from the contention that there is, again within a broadly liberal framework, a common 'presumption against anarchy'; liberals have either prudential or moral reasons to value political stability.

The rest of the work, starting with Chapter 1, is organised around these two sets of reasons. I will begin by considering whether disagreement matters, on a deontological perspective, because its presence acts as a constraint on a state's achievement of legitimacy. The discussion of this theme is going to occupy Chapters 1 to 3, where I will mostly engage with political liberal scholarship. Their conclusion is going to be broadly sceptical regarding the possibility of a fully satisfactory reconciliation between legitimacy and disagreement. If disagreement is taken seriously, or at least as seriously as political liberals generally advocate, then indeed legitimacy becomes a requirement that almost no political order can achieve.

Chapter 1 will give an initial presentation of legitimacy and attempt to develop an account of it. I will observe that the concept of legitimacy in political theory requires at least two different accounts or conceptions, one applicable to the domestic, and the other to the global sphere. In the global sphere, a good account of political legitimacy must refer to a state's territorial sovereignty, that is to say, to the right of a state to the exclusive exercise of authority over a distinct territory. But this presentation of legitimacy presupposes an entire institutional structure already in place; it depends, in other words, on the commonly accepted convention that states are indeed sovereign and autonomous in their exercise of political authority. In this sense, legitimacy understood as territorial sovereignty is a principle that depends, in both its ground and justification, on the acceptance of the conventional interpre-

tation of states as territorially independent entities. Because of its conventional nature, the account of legitimacy that is applicable to the global sphere cannot be affected by the phenomenon of *reasonable* disagreement. Being this the case, then, legitimacy at the global level cannot be what prompts political liberals to consider disagreement as a constraint on political legitimacy.

Chapters 2 and 3 will then discuss legitimacy at the domestic level, and, more specifically, will question the view of legitimacy as public justification that constitutes the main theoretical contribution of political liberalism. The view holds that a legitimate authority is only one that can be showed to be acceptable to each reasonable citizen. I will then proceed to analyse the two different notions of public justification and reasonableness separately in Chapters 2 and 3.

In Chapter 2, I will distinguish the two models of public justification as ‘convergence’ and as ‘reasonable consensus’ and take issue with each of them. ‘Convergence’ fails because it generates excessively demanding requirements on state agents that intend to demonstrate the justifiability of public norms. What is worse, these requirements – namely, that norms are publicly justified only if they are not met with ‘intelligible’ or ‘sincere’ opposition – seem to take for granted that public justification constitutes a self-evidently genuine moral concern, which, is, however, not necessarily the case. The value of public justification for advocates of the ‘convergence’ model seems to depend, more or less implicitly, on a prior unsubstantiated commitment to a libertarian principle of non-interference, which is not, however, subject to the public justification test, thus creating a significant tension within the model itself.

‘Consensus’, on the other hand, fails because of its inability to capture the phenomenon of disagreement in its nuances. Instead of reconciling disagreement

and legitimacy, I will argue, a consensus of reasonable doctrines on principles of justice generates its own form of 'interpretive disagreement'. Reasonable citizens agreeing on the acceptability of certain general principles of justice will still disagree about whether a correct interpretation of such principles can be used to justify particular norms and policies. But this 'interpretive' disagreement cannot be dismissed as inherently less disruptive to the public order than the higher-order disagreement regarding the acceptability of the general principles.

Chapter 3 will focus on the notion of reasonableness, which acts, in the political liberal model of legitimacy as public justification, as the quality or virtue citizens need to possess to be able to reach a consensus on fundamental principles of justice. I will try to elaborate on the notion of reasonableness more in detail, especially in terms of its epistemological underpinnings and implications. I will conclude that reasonableness demands, on the part of the agents that possess it, an attitude of broad scepticism regarding the conceptions of the good they entertain. But, I will argue, a disagreement among agents who are sceptical about the overall credibility of their own conception of the good is already a form of sugar-coated or softened disagreement which, in itself, cannot produce a particularly interesting challenge to the legitimacy of the liberal order.

Chapters 1 to 3 have led to mostly negative conclusions. In those, I have argued that the reconciliation between disagreement and legitimacy is much harder to obtain than the liberal political literature might indicate. Chapter 4 is the first where I will lay out a positive thesis, specifically regarding what equal respect requires in circumstances of political deliberation among disagreeing moral peers. I will argue that equal respect mandates to adopt a positive disposition towards the full comprehension of another's view. Through this cumbersome label, I mean that parties to a political debate who want to treat each other with equal respect ought

to adopt a specific attitude, the willingness to fully understand each other's central point in the discussion. I will further notice how the grounds of this normative requirement lie in an attribution to others of a particular moral status, which I will define 'as-if rationality'. The very recognition of others as moral peers amounts to an assumption (hence the 'as-if' part of the label) that they are as capable as we are to form and revise an intelligible conception of the good. I will further use the chapter to dismiss an idea that has acquired popularity in the literature on pluralism, which is that some of the conceptions of the good existing in society are entitled to a form of official recognition on the part of the state, which should guarantee their equal chance of survival and prosperity.

Chapter 5 will develop the teleological considerations of stability. In working out a liberal theory of stability, however, I will reach conclusions that are significantly deontological in character. I will notice how both political liberalism, with its expectation that the same mechanism used to guarantee political legitimacy (public justification) will also ensure stability ('for good reason'), and political realism, with its attempt to keep stability firmly separate from either legitimacy or justice, fail to provide a satisfactory account of stability. My proposal will focus instead, in a fashion altogether similar to the previous chapter, to the type of attitude that denizens of a political community must display for that community to remain stably well-ordered. I will argue that citizens of a stable liberal society need to adopt an attitude of toleration towards the proposals advanced by their fellows in the public sphere. More specifically, the tolerant attitude I refer entails another form of attribution, parallel to the 'as-if rationality' of the previous chapter. It is the attribution to fellow citizens of a particular moral status, namely the entitlement to try to shape society according to criteria they consider reasonable. Absent that disposition, citizens

would have no motivation to keep complying with norms that clash with their values or interests.

Chapter 6 concludes the work and is where some of the theoretical conclusions I have reached in the previous parts of the book find a practical application. My aim in the chapter is, as common throughout the work, both critical and normative. The practical issue I will dedicate my attention to concerns the conflict between individual commitments and public norms, which many political and legal theorists, along with public opinion, reconstruct as a scenario where some people's integrity is compromised. I will criticise this reconstruction which, I will show, does not offer any valid reason as to why people whose individual commitments are possibly frustrated would deserve any form of accommodation (in the form, for example, of an exemption from the law or an allowance for conscientious objection). I will observe how, by contrast, a reliance on the two core liberal values of equal respect and toleration helps provide a more stable moral basis for the very same policies of accommodations.

I will also point out, however, that this shift from integrity to fairness and toleration leads to important normative revisions. Indeed, moral justifications for accommodations based on integrity tend to emphasise the subjective relevance of the individual commitments, but in so doing are forced into a metaphysically complex and ethically problematic evaluation of the 'depth' of each person's commitment. My account, by contrast, merely starts from the assumptions that people equally possess a capacity and willingness to live according to a conception of the good they have autonomously chosen. This leads to a significantly different treatment of private commitments in the public sphere. Instead of enquiring over the

depth of the commitments that citizens are advancing when they require an accommodation, the state should simply consider whether those citizens can legitimately appeal to the two overarching principles of fairness and toleration.

The Closing Remarks will recapitulate the main argument and better illustrate the relationship between the present work and political liberalism.

0.2 Why Disagreement? Which Disagreement?

0.2.1 The Question of Public Justification in Political and Justificatory Liberalism

Jeremy Waldron identified as the central tenet of liberalism the demand that the 'social order should in principle be capable of explaining itself at the tribunal of each person's understanding' (Waldron, 1987, p. 149). Liberals are united, under this reconstruction, not so much by the particular values they defend, nor by a particular interpretation of these values, but by the demand that whatever hierarchical structure exists in society, it must be justified in terms that all those subject to it may fully comprehend. The 'consent of the governed' of the early social contract tradition might remain no more than a metaphor or a chimaera, but the truly salient requirement of liberalism for Waldron is that the public authority must 'either be made acceptable or be capable of being made acceptable' (p. 128) to every individual.

To this demand for justification, contemporary political liberals respond with a particular view of political legitimacy:³ legitimate states are those that are 'capable of being made acceptable', following Waldron's phrase, to every (adult, autonomous) constituent. The two separate questions of political legitimacy and public justification find then a unitary treatment, and hopefully a unitary solution. In Thomas Nagel's words, the question of political legitimacy consists indeed of 'the history of

³I prefer to reconstruct this way the relationship between justification and legitimacy within political liberalism, rather than claiming, as A. John Simmons does (Simmons, 1999, pp. 756-60), that political liberals conflate the two distinct questions of justification and legitimacy.

attempts to discover a way of justifying coercively imposed political and social institutions to the people who have to live under them, and at the same time to discover what those institutions must be like if such justification is to be possible' (Nagel, 1987, p. 218).

Troubles emerge when one sees that societies contain not only a plurality of individuals but also a plurality of 'tribunals of understanding'. Is the legitimate state then expected to justify its authority before each of the 'tribunals of understanding' that happens to exist in society? That seems an excessively demanding requirement; some of these tribunals might be utterly inaccessible to the uninitiated and others so prejudiced against basic liberal values that no state that commits to them might ever aspire to look justified to them. Luckily, Waldron's justification demand does not seem to commit to the view that the state must always be *justified* to all citizens, but only that it must be *justifiable*.

For political liberals, a model of justification is public, and therefore capable of legitimising the authority of the state, only if it refers to reasons that are indeed shared by, or easily accessible to, the political community at large. These reflections culminate in John Rawls's principle of liberal legitimacy, so presented in *Political Liberalism* (hereafter, PL):

[O]ur exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason. (PL, 137)

What does it mean, however, that the essentials of the constitutions must be such that all citizens 'may reasonably be expected to endorse them'? As I will dedicate two chapters to discussing the best way to interpret this phrase, I should now

simply lay down some preliminary observations, which mainly aim at introducing the literature I will engage with in the first half of the work.

One particular way to interpret what it means for constitutional essentials to be such that all citizens, as free and equal, must reasonably be expected to endorse them is the one we find in the classic political liberalism of John Rawls himself (PL), Thomas Scanlon (Scanlon, 2003), and Thomas Nagel (Nagel, 1987), later defended by, among others, Samuel Scheffler (Scheffler, 1994), Charles Larmore (Larmore, 1996; Larmore, 1999), Stephen Macedo (Macedo, 1997), and Jonathan Quong (Quong, 2011). A core notion within this current of thought is ‘reasonableness’, representing the idealised standard that the different doctrines and non-political conceptions endorsed by citizens in society (which Rawls collectively defines as ‘comprehensive doctrines’)⁴, have to meet, in order for an agreement on the acceptability of certain fundamental principles of justice to be possible. Within a single society, Rawls argues, reasonable doctrines will tend to overlap on a common political conception that includes fundamental principles of justice. The ‘overlapping consensus’ that is thus construed (Rawls, 1999 [1987]) (PL, pp. 133-172) acts as the basis of public justification and, following the aforementioned liberal principle of legitimacy, of the legitimacy of the state; when a consensus of this form comes about, individual citizens see, through reasons they possess and that are part of their own conceptions

⁴The best definition of comprehensive doctrines is in PL, p. 58, where Rawls is in fact defining ‘reasonable’ comprehensive doctrines (but all that he says there seems to have to do much more with comprehensiveness than with reasonableness). He recognises three features of a (reasonable) comprehensive doctrine. Firstly, a comprehensive doctrine is an ‘exercise of theoretical reason [as] it covers the major religious, philosophical, and moral aspects of human life in a more or less consistent and coherent manner’. Secondly, ‘it organises and characterises recognised values so that they are compatible with one another and express an intelligible view of the world’. Thirdly, ‘while a reasonable comprehensive view is not necessarily fixed and unchanging, it normally belongs to, or draws upon, a tradition of thought and doctrine’. Rawls insists probably more than necessary on the comprehensiveness and internal coherence of the doctrines he believes together compose the public debate; as it is going to be evident starting from Chapter 4, I will take a more minimalist definition of what counts as a citizen’s conception of the good life.

of the good, that the authority they are subject to is a legitimate source of constraints.

I will dedicate an entire chapter to the concept of reasonableness, but let's introduce some preliminary notes here. Thomas Scanlon, while elaborating his principle of 'reasonable non-rejectability',⁵ describes as unreasonable the attitude of those agents who are unwilling to reach an 'informed, unforced, general agreement' (Scanlon, 2003 a, p. 132) with other parties. Rawls, on the other hand, introduced two conditions of reasonableness in *Political Liberalism*: the willingness to propose 'fair terms of cooperation' (as opposed to terms of cooperation that simply benefit oneself) for the society in which one lives and the reconnection of the origins of pluralism to the 'burdens of judgement' (PL, pp. 49-50, 61). The 'burdens of judgement' are in turn identified as those natural impediments in the free use of the faculty of judgements that prevent a condition of reasonable agreement from becoming stable.⁶ Hence, in recognising the 'burdens of judgement' as the ultimate source of pluralism, reasonable agents interpret disagreement as the most natural outcome of a reasoning process under conditions of non-coercion. To summarise, reasonable people are those that 1) are happy to propose terms of regulation for the society in which they live that they expect others can reasonably accept and 2) in case of disagreement, consider others' position as, albeit wrong, possibly grounded in a reasoning process that is as valid in principle as the one that has grounded their own views.

⁵ 'An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for informed, unforced general agreement.' (Scanlon, 2003 a, p. 132)

⁶ 'An explanation of the right kind [about the cause of reasonable pluralism] is that the sources of reasonable disagreement – the burdens of judgement – among reasonable persons are the many hazards involved in the correct (and conscientious) exercise of our powers of reason and judgement in the ordinary course of political life.' (PL, pp. 55-56)

The intellectual focus on public justification also grounds the interest in disagreement in the other major strand of liberalism. Within the so-called ‘justificatory liberalism’ defended by Gerald Gaus and his followers,⁷ disagreement acts as the background against which a political authority is bound to attain legitimacy. What characterises ‘justificatory liberalism’ and distinguishes it from ‘political liberalism’, however, is the avoidance of the idealisation that underlies the concept of reasonableness; this implies that public justification cannot simply appeal to reasons that citizens would possess under certain, reasonably idealised, conditions. The public authority’s legitimacy will depend, on the contrary, on its being accepted through reasons citizens *actually* possess, even though (and this further point partly re-introduces some form of idealisations) they may not be currently aware of possessing them. A reason (which we can simply consider, following Scanlon, as a consideration in favour of an action)⁸ can be possessed by an agent without her being aware of it in case the agent has failed to consider the full implications of her own beliefs; to exemplify, if I believe both that human life begins at the moment of conception and that murder corresponds to taking human life, I then possess a (*pro tanto*) reason to oppose abortion, whether or not I am actually aware of it.⁹

Gaus holds that a norm or institution¹⁰ is publicly justified when it is ‘not open to reasonable doubt’ (Gaus, 1999, p. 276), which means that, in the aftermath

⁷ See (Gaus, 1996) (D’Agostino, 1996) (Gaus, 1999) and (Gaus, 2011 b) and, for recent defences, (Vallier, 2011) (Van Schoelandt, 2015).

⁸ See for example (Scanlon, 2006, p. 231).

⁹ Gaus believes that through this model he has overcome the traditional ‘internalist’/‘externalist’ dichotomy in the debate about motivational reasons. Indeed, considering agents’ reasons as those they possess once they have reasonably considered the full implication of their beliefs does not reduce this view to motivational ‘internalism’, insofar as the reasons that matter in public justification are derived from but not identical with the agents’ actual beliefs (in Bernard Williams’ terminology, their ‘subjective motivational set’ (Williams, 1981 a, p. 102)).

¹⁰ The exact scope of public justification is also debated, with ‘justificatory liberalism’ advocates generally adopting a less restricted view of the appropriate scope of public justification. I will explore this point further in the next chapters.

of a discursive process,¹¹ it is universally accepted by all the participants to the discussion, provided they are at least committed to the full analysis of their own ideas. As this requirement is too stringent in most occasions, especially due to the circumstances of reasonable pluralism, Gaus accepts that another manner in which public authority can be justified is in its role as an ‘arbitrator’ or ‘umpire’ capable of breaking the indeterminacy caused by disagreement.¹²

A further challenge for ‘justificatory liberalism’ is that, for an institution or a norm to be publicly justified, its justification must appeal to reasons that are, by definition, ‘public’, that is, in some sense compelling and understandable for the body politic at large. Nothing in the account presented so far implies that the reason must be unique; public justification can appeal instead to a plurality of reasons, each compelling to a subset of the constituency.¹³ However, the further requirement that these reasons must be understandable to the constituency at large leads to the task of identifying which is the plausible account of ‘shared understandability’ to be employed. Gaus’s ‘principle of sincerity’¹⁴ was the first attempt within the ‘justificatory liberalism’ paradigm to respond to this challenge. Kevin Vallier recently proposed a partly different test of ‘intelligibility’ that should be capable of overcoming the limits of idealised ‘accessibility’. According to Vallier, ‘A’s reason X is intelligible to the public if and only if members of the public (at the right level of idealization) can see that X is justified for A according to A’s evaluative standards’ (Vallier, 2011, p. 388).

¹¹ Gaus is influenced in this respect by Habermas’s ‘discursive ethics’ (Habermas, 1996).

¹² ‘The crucial task of government, then, is to serve as an umpire, judge or arbitrator, providing a practical resolution of our reasonable disputes about justice’ (Gaus, 1999, p. 279).

¹³ Fred D’Agostino (D’Agostino, 1996) conceptualises this distinction between two possible ways of agreeing on something as a distinction between *consensus* and *convergence*. I will rely a lot on this distinction in §2.1 and §5.2.

¹⁴ ‘Betty’s argument justifying N to Alf is sincere if and only if (1) she is justified in accepting N; (2) she has a justified belief that N is justifiable in Alf’s system of reasons and beliefs’ (Gaus, 1996, p. 141).

The presentation of ‘justificatory liberalism’, albeit too brief to ground a general assessment for such a complex theory, was functional to understanding, in particular, the differences and similarities with classic political liberalism. In this regard, it is worth noticing how, within justificatory liberalism, the focus on public justification seems to be grounded in a general principle of ‘presumption in favour of liberty’,¹⁵ according to which any interference on others’ freedom creates a burden of justification that the interferer has to discharge. This is a familiar principle in liberal philosophy, but, it must be remarked, not identical, despite some similarities, to the principle of legitimacy through public justification mentioned before, mainly because the latter only concerns coercion exercised on the part of a public authority.

One consequence of this distinction is that, because of their ultimate interest in the justification of interferences, ‘justificatory’ liberals cannot accept as a solution to the problem of disagreement the idealised ‘overlapping consensus’ of political liberalism. Whether the exclusive focus on reasonable doctrines is interpreted as a result of an idealisation of democratic societies (arguably, the most plausible reading of Rawls) or as an effective restriction of the pool of citizens who are entitled to public justification (as in Jonathan Quong)¹⁶, the consensus of reasonable views cannot work, according to Gaus and his followers, as a comprehensive account capable of justifying publicly any possible breach of freedom. Indeed, the limited scope of the

¹⁵ (Vallier, 2018) does not make my distinction and consider the principle as ultimately supported by Rawls himself. Of the authors that explicitly advance the principle of ‘presumption in favour of liberty’, it is worth mentioning in particular, for the influence his philosophy had on Gaus, Stanley Benn, who defended the principle in (Benn & Weinstein, 1971).

¹⁶ In his important reformulation of the Rawlsian proposal, Quong upturns the chronological relation between the overlapping consensus and the achievement of a democratic well-ordered society. For Quong, the achievement of a consensus on ‘the more fundamental idea of society as a fair system of social cooperation between free and equal citizens’ (Quong, 2011, p. 162) is a pre-requisite for having any discussion about public justification. This way, citizens who have not reached a consensus on such fundamental values are excluded from any scheme of public justification that the public authority can offer.

political consensus, which only includes for Rawls matters of constitutional rights and regulations about the ‘basic structure’ of one’s society,¹⁷ is incapable of overcoming the problem of everyday coercion.

0.2.2 Respect, Vulnerability and Stability: Why Liberal Theories of Justice Cannot Ignore Disagreement

We have seen that, for both political and ‘justificatory’ liberals, disagreement matters as a fundamental challenge to the state’s achievement of public justification and, by implication, legitimacy. As I am going to criticise this take on why disagreement is fundamentally interesting within a (liberal) theory of justice, I want to provide an alternative explanation. I suggest that what matters for a disagreement to be relevant for justice is that one of two conditions obtain; the disagreement must be either (1) of particular import to some agents’ conscience or (2) capable of constituting a genuine threat to political stability. I will define the first as the *deontological* and the second as the *teleological* reasons for why disagreement matters in a liberal theory of justice.

Why does it matter that some value-disagreements are of import to some people’s conscience, whatever we mean with this ambiguous term?¹⁸ I defined ‘value’ before, in the most ecumenical way I can imagine, as whatever is taken to be relevant, beyond the mere strive for survival, in human life. The centrality of subjective values to individual lives can easily become, however, a source of vulnerability; people are routinely discriminated, marginalised, stigmatised for the values they have chosen to adopt as meaningful to their life. Yet, the circumstances of public deliberation sometimes demand from individuals that they disclose publicly a part

¹⁷ More on Rawls’s restrictions on public justification in §2.1, 2.3.2 and 5.1.2.

¹⁸ For an attempt at formulating a liberal theory of conscience, see (Brownlee, 2012).

of their identity that can be the object of criticism, ridicule and other stigmatising attitudes.¹⁹

Take two kinds of societies where this demand for exposure does not come about. The first is a completely homogeneous society, whose members all tend to embrace a single set of perfectly compatible values and an identical ranking mechanism for when said values have to be weighed against each other. The second is a multicultural authoritarian society which is made up of various communities that have complete control over their members' expression of a conception of the good; right to exit is never recognised. In the first type of society, exposure of one's conception of the good would be unproblematic because people would express something that is recognised by their fellows as a familiar trait that they would never feel inclined to criticise. In the second type of society, people's public identity is imposed on them; members of the various communities have no choice on which value of theirs they can publicly expose or conceal. In fact, their public exposure is unnecessary, as they are already classified, without their consent, as faithful acolytes of particular conceptions of the good, with all the values attached to these.

Neither of these societies is, I am sure, particularly attractive; I imagine most people would find the second society dystopian and the first, more trivially, just a boring place where to live. Nonetheless, in both societies citizens are shielded from that particular vulnerability that consists in having to bring to the public something that is at the same time central to one's identity and extremely controversial outside of it. They never have to take a defensive stance, that is, towards their own values and ideas of the good.

¹⁹ For the relevance of privacy and concealment to a liberal theory of autonomy, see especially (Nagel, 1998) (Marmor, 2015).

By contrast, in societies characterised by a significantly rich public forum and by endemic and irreducible value-disagreement, people are sometimes required not only to disclose the values that are relevant to them but also to explain to their fellow citizens why these values ought to receive some kind of political recognition. I do not want to exaggerate this point; democratic politics does not require a constant and direct public involvement of every single citizen. In fact, one of the reasons why we value so much the secrecy of the vote for citizens, and by contrast we do not ordinarily allow elected representatives to vote in secret in public assemblies, is presumably because we want citizens to give a contribution to politics without paying any unnecessary costs in terms of personal involvement. Still, in order for citizens to exercise their vote meaningfully, some ideas must enter the public sphere; ideas that will contain in many cases values directly derived from some citizens' conceptions of the good. At least some elected representatives, therefore, must be capable of introducing to the public sphere not just the preferences and interests of their trustees, but also their values.

What I want to point out through these – admittedly still vague – reflections is the idea that the combination of public deliberation and value-disagreement can place inhabitants of contemporary democratic states in an uncomfortable position, where they are the object of disrespect, marginalisation, or ridicule because of ideas of the good they have chosen to adopt and that they can never completely conceal from others. This condition of persistent vulnerability is the main source of what I called before the deontological reasons to take into account the existence of value-disagreement in a liberal theory of justice. As we know that this vulnerability exists, we have deontological reasons to prevent the emergence of a situation where some citizens, for example because more capable of commanding power or resources, can exploit it to marginalise their fellows even further and reinforce their condition of

superiority. I am going to suggest in my Chapters 4 and 5 that we can do that by letting a particular culture emerge in our society – a culture, specifically, of toleration and equal respect.

As the previous reflections, due to their vagueness, might not be decisively compelling, I want to offer another reason to take into account the fact of disagreement. The reason resides in the value of political stability, derived from the belief that the public authority is necessary to guarantee the conditions in which human lives can flourish in their diversity without interference. This is an ideal, if possible, even more primitive in liberal morality than the one of legitimacy through public justification which, we have seen, political liberals take to be foundational. For authors who are close to some form of multicultural libertarianism, this is indeed the only ideal liberals can consistently advance if they want to be sensitive to the differences that exist in society.²⁰ Kantians, starting from a different perspective, agree in placing importance in political stability, insofar as they consider the state as the agency necessary to enforce a system of rights capable of guaranteeing individuals' moral independence.²¹

²⁰ See especially (Galston, 1995): 'properly understood, liberalism is about the protection of diversity, not the valorisation of choice. [...] To place an ideal of autonomous choice – let alone cosmopolitan bricolage – at the core of liberalism is in fact to narrow the range of possibilities available within liberal societies' (p. 523). Although disagreeing with Galston's statist position, Chandran Kukathas also offers a defence of liberalism as a theory centrally about the toleration of various and competing conceptions of the good life in (Kukathas, 2003). What distinguishes Kukathas's proposal is the interpretation of the liberal state as an 'archipelago' of free associations, a 'society of societies' (p. 22) each operating under the principle of freedom of association and holding a certain amount of authority.

²¹ 'We are not really independent from one another if we cannot rely on our rights. [...] Only a state that protects the familiar liberal rights and democratic procedures will do. And if this is right, we need to be in a liberal state in order to be able to treat each other as independent' (Pallikkathayil, 2016, pp. 176-177). Similar reflections are in (Waldron, 1996) (Ripstein, 2004) (Hodgson, 2010) (Pallikkathayil, 2010) (Flikschuh, 2010) (Sinclair, 2018).

There is, to simplify, what we can call a ‘presumption against anarchy’ within liberal literature that notoriously dates back, although with different sorts of arguments, at least to Hobbes and Locke.²² It can be fleshed out by saying that, although liberals place a significant focus on the limits of the activity of the state due to individual rights, it is nonetheless necessary to have an agency that guarantees those very same rights. Even liberal theorists who subscribe to a form of anarchism whereby no political obligations can be justified,²³ have to accept that the liberal state is prudentially valuable, insofar as it is the institution that is less than any other conducive to interferences with personal autonomy. It follows that all liberals accept that political stability, which I will now preliminarily define as the continuation in time of the authority of the state, is at least instrumentally valuable.

Disagreement constitutes a potential threat to political stability, because citizens who deeply disagree with a norm may possess a subjectively compelling reason to perform a series of acts or omissions that can globally compromise political authority. These acts or omissions vary from the manifestation of dissent before or after the enactment, to the refusal to comply with a law that is considered unjust, to the complete withdrawal of political allegiance in the most extreme cases.

Identifying the reasons that may prompt scholars interested in designing a liberal theory of justice to be interested in the theme of disagreement was my main task in the present section. Now, before I come to the substantive part of the work, a final preliminary task will consist in fleshing out the ambitions and limitations of

²² Specifically, Hobbes’s rejection of anarchism is primarily prudential and (proto)-utilitarian: life in the state of nature being ‘nasty, brutish and short’, there is no alternative for self-interested agents that to leave that grim scenario and subject themselves to the will of the sovereign. For Locke, by contrast, who anticipates Kant in this regard, the justification for having ‘civil society’ relies primarily on getting rid of the arbitrariness that characterised the enforcement of the natural law and the respect of natural rights within the state of nature.

²³ (Wolff, 1998 [1970]); (Simmons, 1979).

the entire work. I will thus illustrate in the next section both what this work aims to achieve and how it came about.

0.3 Limits and Aspirations

0.3.1 In Defence of (Ideal) Theory

The point of this section is essentially apologetic. I aim to anticipate some of the objections regarding the structure I gave to the investigation, informing the reader on crucial features that have characterised both my initial interest in the question and its treatment in the present work. Hence the title of the section, which refers to both the objectives I did and did not set out for myself in the present work and to the general ambition I cultivated in trying to meet them.

What am I therefore aiming at, by looking for a solution to the question of disagreement in politics? I will firstly give a series of negative answers through which the methodological principles informing this work will be disclosed.

The solution I will delineate is normative and political, but not in the sense that it can lead to immediate policy recommendations. In the final chapter, I will discuss some existing policies, and assess their broad viability in the light of my normative conclusions. I will not advance, however, a list of recommendations regarding the policies that the fact of disagreement demands us to adopt. The reasons for this do not merely reside in the limited pages of the work but in the deeper methodological limits of the discipline.

Normative political theory, as I conceive it, does not work, if not in the manner of pure examples, with the empirical data that compose the appropriate context of political practice. The choice of policies is a task that necessitates careful consideration of those data; the possibility of success very much depends on them. These

data that policy-makers need to consider belong to an almost infinite array of categories; they are such things as citizens' day-to-day preferences, cultural and historical factors, general constraints on institutional design, natural circumstances, and so on.²⁴

A political *theory* cannot accommodate all these infinite contextual data; they are too many and too various to be captured under a unifying pattern that a theory must necessarily possess. Or at least, it cannot accommodate all these data and still produce something that has even minimally universalistic aspirations.

But maybe, universality can be sacrificed. Advocates of renewed 'political realism' in political theory would claim that, because political solutions require attention to context, and universalistic political theory necessarily abstracts from context, universality is expendable. If allegedly universalistic political theory, in other words, is incapable of producing solutions to concrete political problems, so the worse for universality.²⁵

²⁴ A reader can easily recognise the broadly Hayekian tone of these reflections, especially in the brief description I gave of contextual data, which is influenced by Hayek's reference to the 'particular circumstances of time and place' in (Hayek, 1945, p. 522). Hayek would be sceptical about the possibility itself that policy-makers might ever accumulate a sufficient amount of what I have called contextual data to produce effective policies, but this type of scepticism is beyond the scope of the present work. A scepticism that is instead germane to my reflections is the one Hayek expressed about economic theory, and its hopeless attempt to advance predictions without complete familiarity with such contingent, contextual data, as in his Nobel memorial lecture (Hayek, 1989).

²⁵ See for example this statement from two authors: 'Contextualism [...] provides a way into the rejection of ambitious attempts to formulate grand theories from which to derive prescriptions for any possible scenario [...] in favour of a normativity that is appropriately sensitive to the specific conditions under which political decisions are taken and agents act' (Rossi & Sleat, 2014, p. 694). Or, even more clearly: 'The universalist tendencies of contemporary political philosophy and theory are part of what prevents it from addressing real political situations, which always involve actual political actors with particular disagreements (Jubb, 2017, p. 116).

The challenge I have set myself in the present work is that of salvaging universality, without that entailing the opposite conclusion, 'so the worse for real politics'. How can this be done, and why do I believe that there is a value in salvaging universality?

My acceptance of universality needs refinement. My first reason for rejecting a mainly contextualist conception of political theory is a form of scepticism. I am unpersuaded by the belief that rejecting universalism and taking context into account can make normative political theory a discipline capable of producing immediately action-guiding recommendations for political practice. As I claimed before, which norms ought to be implemented in a specific society depends on so many contextual contingencies that a theory could never efficiently accommodate. Crudely put, those whose aspirations are in the elaboration of norms and policies that take into account, as much as it is possible within the limits of human cognition, the highest amount of relevant contextual data, immediately put themselves out of the field of political theory, entering instead that of public policy or legal elaboration.

A certain amount of idealisation is inherent to the very existence of a theory, in the basic sense that a theory, as opposed to a mere list of unconnected recommendations, necessarily abstracts from at least some features of reality, and this is true in the scientific as well as in the philosophical fields. Indeed, we speak of a 'theory of x' (for example, the theory of evolution) to isolate the area of reality that, in the scientific realm, is captured by the theory itself. As events in the real world are affected by a plurality of causes, why things in the world are in a certain way will always be explained only partly by a particular 'theory of x'; all the rest is abstracted away.

It follows that the expression 'non-ideal theory', popularised in the recent debate following a statement in Rawls's *Theory of Justice* (hereafter, TOJ), is, if taken

literally, an oxymoron²⁶ and must be taken to mean something slightly different, namely a less idealised theory, or, following from before, a theory that takes more contextual data into account. In Rawls, for example, the ‘contextual data’ that are taken into account in what he deems ‘non-ideal theory’ concern ‘natural limitations and historical contingencies’ (TOJ, 216) and the fact of ‘partial compliance’ with norms in society (TOJ, 217). Both are facts of the world that, according to Rawls, a theorist of justice can legitimately set aside at the level of ‘ideal theory’, insofar as the objective is that of elaborating a ‘conception of a just society that we are to achieve if we can’ (TOJ, 216). As these facts, presumably among many others, will prevent us from achieving the just society designed by the principles of ideal theory, such principles will have to be slightly readjusted when more facts are gradually taken into account, precisely because a strict adherence to the principles of ideal theory will not ensure that justice comes about in all possible circumstances.

The objectives set out by ideal theory, however, remain the same throughout; in fact, following A. John Simmons’s reconstruction of Rawls, we can say that ‘[n]on-ideal theory requires us to work toward achieving the perfectly just basic structure that *would* result from strict compliance with the *ideal* theory of justice’ (Simmons, 2010, p. 18). Which immediately sparks the question of why we then need an ideal theory in the first place, since we cannot directly derive from it normative guidelines that we need to follow, here and now, to eventually reach justice. Clearly, someone might object, we do not need an ideal of justice to realise that certain conditions of

²⁶ Of course, depending on how one defines ideality, the opposite can also be true, namely, that because a theory of justice is one that needs to deal with some forms of imperfection, for example human limited altruism, then a completely ideal theory of justice is nonsensical. But this objection to ideal theorising, defended for example in (Levy, 2017), has no bite, because no allegedly ‘ideal’ theory of justice abstracts away from such circumstances of human nature, such as humans’ limited altruism, that would render a theory of justice itself meaningless.

the existing world need, for reasons of justice, to be remedied.²⁷ Hence, why not arguing directly from the circumstances of injustice and see what is the most appropriate strategy to remedy them?

I am sceptical that we could call a simple list of recommendations regarding the gradual alteration of a certain state of affairs, derived from the intuitive realisation that the state of affairs in question is unjust, as a theory. But here is where the bulk of the problem ultimately is, because someone could well bite the bullet and reply that we do not need then a theory of justice in the first place, whether ideal or non-ideal. I will now offer a response to this position, which is hardly original but, I believe, often neglected by the critics.²⁸

Justice is an essentially complex concept, for which I mean one that never denotes the possession of a single property in the real world. Physical complexities about the perception of colour aside, the concept of 'redness' purely denotes the entities in the real world that possess the property of being (or, more correctly, being perceived by the human eye under certain conditions as) red. By contrast, a just or

²⁷ This insight, which is of course correct, is central to a plethora of critiques to ideal theorising in politics, all of which are, if my reasoning is sound, slight misinterpretations of what ideal theory actually consists in, because they tend to conceive the possibility, which I consider completely paradoxical, of a political theory that is deprived of any form of idealisations. See in particular (Sen, 2006) (Schmidtz, 2011) (Mills, 2017).

I further disagree with the presentation of ideal and non-ideal theory presented in (Hamlin & Stemplowska, 2012). Alan Hamlin and Zofia Stemplowska introduce a distinction between a 'theory of ideals' whose 'purpose is to identify, elucidate and clarify the nature of an ideal or ideals' and 'ideal-non-ideal theory', which deals with 'the social arrangements that will promote, instantiate, honour or otherwise deliver on the relevant ideals' (p. 53). As I see it, this distinction is superfluous. An ideal of justice is, in political philosophy, indistinguishable from a social arrangement that is presumed to realise justice. Taking Rawls as an example, what difference is there between the principles of justice, which are certainly about the arrangement of an ideally just society, and an 'ideal' of justice? One could maybe try to give a purely conceptual, abstract conception of justice, which abstracts from what a just society looks like. Whatever the merits of that exercise, they lie outside political philosophy.

²⁸ There are many similar defences of (abstract, ideal) theory in ethics and politics. Apart from Rawls's reflections in chapters 7-9 of *Theory*, two of the most compelling (and sources of inspiration for these paragraphs) are (O'Neill, 1987) and (Hooker, 2012).

unjust state of affairs is one that is analysable in terms of, and possibly reducible to, the possession or non-possession of other moral qualities. If I call a situation unjust, it might not be obvious at first sight why I hold that belief, but, if I am asked to justify it, I will surely rely on further moral categories that explain exactly which components of justice have been, in that scenario, violated. So, I might say, for example, that the gender pay gap is unjust *because* (once I'm asked to justify it) it violates gender equality. I certainly do not need a complete theory of justice to hold the intuitive belief that the gender pay gap is unjust, and I might not need one even to justify that belief in terms of violations of an equality principle.

But now suppose I have in front of me a pair of possible solutions to remedy the issue of the gender pay gap, which are so distinguished. Solution A promises to address the problem swiftly, but its implementation will probably lead to sacrificing, to a significant extent, another fundamental ingredient in the complex concept of justice, say individuals' freedom. Solution B works at a slower pace, but its costs in terms of restrictions on individuals' freedoms are negligible.

How do we operate a selection of one of these options? The problem we encounter in taking this decision is one of weighing different moral concerns and their respective urgency and dispensability. Following Rawls (TOJ, 36-40), we define this as the 'priority problem' in a theory of justice. Some moral pluralists, usually of an intuitivist persuasion, believe that moral theory purely amounts to a kind of systematisation of our intuitive beliefs about the right and the good, the outcome of which can only be a list of *prima facie* motivating moral concerns.²⁹ But political theory needs something more than that, namely, some indications, even of a purely provi-

²⁹ For a classic defence of this position, see (Ross, 2002 [1930]), and, for a recent resurgence, (Dancy, 1993).

sional and revisable kind, about which kinds of *prima facie* motivating moral concerns ought to be prioritised in case of conflict. Of course, it can well be the case that many cases constitute authentic moral dilemmas, but that cannot be a conclusion we want to universalise for every case where more than one moral concerns are at stake. Otherwise, we would be led to the conclusion that we are morally entitled, in every situation where a genuine conflict of moral concerns emerges, to simply toss a coin.

Now, the fundamental role of a theory is precisely that of turning that ‘coin-tossing’ into a more rigorous exercise of appropriate balancing of different motivating considerations, both at the individual level (moral theory proper) and at the social one (a theory of justice). ‘More rigorous’ does not entail that the theory must produce a perfect ‘algorithm’³⁰ that one can use to derive from the general principles unambiguous recommendations for every possible situation of conflict. But it does entail that general principles of balancing must exist to solve the ‘priority problem’, precisely to prevent the ‘coin-tossing’ response to conflict from becoming overly common. As Rawls famously said about W.D. Ross’s intuitionist pluralism, ‘[a]n intuitionist conception of justice is, one might say, but half a conception’ (TOJ, 37). Notice that Rawls’s accusation against the pluralist is not that of producing ‘half a theory’, because the pluralist might well be content with that, but ‘half a conception’, which means that, without a solution to the ‘priority problem’, we cannot even start giving an answer to the question of what justice requires.

The previous paragraphs should suffice to explain why, in a work that is devoted to analysing the problem of the reconciliation between value-disagreement and justice in a social environment, I will employ a broadly theoretical approach that

³⁰ (O’Neill 1987) presses this point.

will not shy away from certain forms of idealisations, provided that these are justified from the outcomes they are supposed to produce in terms of intellectual clarity and simplicity. The disadvantages of this approach were already presented; because a theoretical, (mildly) idealistic approach cannot take into account a sufficiently significant number of contextual data, we will not be able to immediately extrapolate out of it normative recommendations which we can apply straight away. This is true, as I argued before, of *any* theory of justice, regardless of whether it purports to be less 'ideal' than the Rawlsian model. In a slogan, political practice that abstracts from the appropriate contextual data is useless, but political theory that aspires at including the sufficient number of contextual data to lead to immediate practical implications is hopeless.

In conclusion, I do not see why ideal theory is not a validly pursuable project. A normative theory without idealisations would be no other than a list of short-term recommendations about the treatment of a particular issue; equivalent, in this sense, to a mere manual of instructions. Idealisations are needed the moment we start conceiving the problems we are treating (such as disagreement in the present work) as only existing within a complex conglomerate of equally relevant problems. Then, we recognise that without some form of idealisation the investigation cannot even commence, because taking into account all the infinite contextual factors that might have an effect on the problem we are investigating would be virtually impossible.

Of course, we will be ready to consider more and more of these factors once the investigation has led to significant results and can now be translated into concrete policies. Until that moment, however, universality must take priority over concreteness, as we want to produce some principles that, albeit not immediately applicable as real-world policies (due to their limited incorporation of contextual

data), can be appreciated as valid and motivating across the most varied contexts and circumstances, provided that these contexts share some basic moral values. This is, in sum, the aspiration I set out for the present work.

1. Political Liberalism and the Question of Political Legitimacy

Although the literature following the publication of *Political Liberalism* sometimes seems to cultivate an interest in public justification as an independent moral concern (together with the cognate ‘public reason’), the ultimate reason why public justification matters within political liberalism is as a solution to the question of political legitimacy. To reiterate, a coercive authority is legitimate according to Rawls (at least in a sense)¹ when ‘it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason’ (Rawls, 2005 [1993], p. 137).

If this is the conceptual connection between legitimacy and public justification, then my previous point about disagreement being a possible obstacle to public justification should be reframed as disagreement being an obstacle, *a fortiori*, to political legitimacy. This use of public justification as a solution to the question of legitimacy is exactly what, however, I plan to criticise in the present and following two chapters.

Before that, however, we need to define what legitimacy means. This, it is worth noting, is a conceptually different question from the one about the solution of the question of legitimacy. If I take public justification to be the solution to the question of political legitimacy, I have not provided any information regarding the definition of legitimacy. I have simply identified a problem (political legitimacy) and furnished a solution (public justification), but I still have provided no clue regarding what the problem in the first place is. Or, to rephrase in yet another way, I have ex-

¹ I am going to elaborate on this point later in the chapter.

pressed the necessity of a logical implication leading from public justification to political legitimacy (namely, if public justification then political legitimacy), but nothing has been proved about legitimacy itself. It might well be that the implication goes both ways, and that, as a consequence, legitimacy and public justification are practically interchangeable notions (as I am going to show later, this is explicitly denied by Rawls), but this is not what the principle of legitimacy through public justification amounts to.

The aim of this chapter will be that of filling the gap created by the introduction of the principle of legitimacy through public justification. It will do that by presenting a possible model of political legitimacy. The next two chapters will then investigate whether the connection between legitimacy and public justification is, at least at the domestic level, as straightforward as the political liberal literature seems to affirm.

The first task of this chapter will be that of discharging the ‘anarchical challenge’, which is one of the most common, but, I aim to show, most implausible, attempts at showing the normative role of political legitimacy. After dismissing that, I will suggest that legitimacy can be differently interpreted in two contexts, domestic and global.

Domestic legitimacy is the key concept discussed in the next chapter. However, and with the main aim of explicating the distinction from legitimacy at the global level, I will already provide in this chapter a possible account of domestic legitimacy, defining legitimacy in the domestic context as the right of a public authority to issue political obligations. For an *account*, I will here mean a possible conception of the ‘contested concept’² of domestic legitimacy, that is to say, an answer to the question ‘what is domestic legitimacy?’. I will leave to the next chapter, on the

²(Gallie, 1956).

other hand, to present a possible *standard* of domestic legitimacy, for which I mean an answer to the different question of which features states in the real world need to possess, in order to be called legitimate.

I will then move the discussion to the global context. My main argument there will be that global legitimacy, when taken to be equivalent to a right to territorial sovereignty, is a primarily conventional normative notion, for which I mean a normative notion whose ground and justification depend to a vast extent on the existence of social practices.

I will reflect on the implications of this conclusion for the relationship between global legitimacy as territorial sovereignty and disagreement. I will observe that, as a primarily conventional normative concept, territorial sovereignty cannot be subject to the phenomenon of reasonable disagreement. Hence, this type of legitimacy cannot be the one political liberals have in mind, when they consider public justification (as justification offered to reasonable citizens) to be a solution to the question of legitimacy.

1.1 A Model of Political Legitimacy

1.1.1 Rejecting the 'Anarchical Challenge' Model

In many works in political philosophy, starting of course with Hobbes' *Leviathan*, the question of political legitimacy has a direct addressee, the anarchist. In considering the institution of the state ultimately unjustified, the fictional anarchist has the same intellectual role of the sceptic in early modern discussions in epistemology; it represents the first stage towards persuasiveness. Once the anarchical alternative has been refuted, and the anarchist has been persuaded to abandon her commitment, it seems that something relevant, albeit embryonal, has already been shown, namely that political legitimacy is at least a meaningful category that can be applied

to some entities in the external world (in the same sense as the Cartesian refutation of the sceptic shows that knowledge of the external world is not always an illusion). This is the argumentative path followed, in modern scholarship, by Robert Nozick (Nozick, 2013 [1974], pp. 3-9), but it is one I consider ultimately unattractive. The reason why I believe this type of reasoning is unattractive is that it must necessarily presuppose, as it is indeed the case for Nozick (inspired in this by Locke), the existence of pre-institutional moral values that can be used to assess morally, by themselves, the legitimacy of institutions.

I will now show why I believe the *exclusive* reliance on pre-institutional values to assess the legitimacy of institutions is problematic. The reason why I take the pre-institutional thesis to be problematic is that in a plausible reading of the principle of legitimacy, the existence of the state turns out to be a pre-condition for any meaningful use of the notion of legitimacy itself. This being the case, legitimacy becomes a notion unfit to sustain the anarchical challenge, as it already presupposes for its meaningful use the existence, actual or hypothetical, of a public authority to be assessed.

Imagine the liberal principle of legitimacy through public justification being rephrased so to acquire this conditional form; '*if* there is an authority requiring compliance, then it must be justified in terms that those subject to it can see as reasonable'. This presentation grounds³ the principle of legitimacy in the existence of institutions, but only in a conditional form; which means that the existence of a public authority does not *imply* the truth of the principle of legitimacy, but renders it *relevant* and *applicable*.

³ I am here using 'grounding' in the epistemological sense of lending support to another principle. I am not using it in the strict Cohenite sense (Cohen, 2003) which, as noticed in David Miller's celebrated critique of Cohen (Miller, 2013), tends to merely reduce grounding to logically implying.

If this were a plausible reading of the principle of legitimacy, Nozick's preoccupation would be misplaced. If the principle of legitimacy had a conditional form that depends, for both its relevance and applicability, on the existence of public authority, then the anarchical objection has already been addressed; the anarchist has been offered a standard through which she can judge whether certain states are legitimate. It might well be that all states are, contingently, illegitimate, in the sense that no existing or existed state meets or has ever met the criteria of public justification, but the conditional statement implies that the logical possibility of a legitimate state cannot be denied. In other words, the conditional reading of the principle of legitimacy implies that there is at least one state in one possible world that is legitimate, namely the state that meets the criteria of legitimacy. As we are not (yet) concerned about showing that there are states in our current world that meet those criteria, or that there could feasibly be some in the near future, the anarchical challenge has already been dismissed by conditioning the notion of legitimacy on the existence of a public authority.

Now, imagine someone replied that the conditional reading of the principle of legitimacy is implausible; legitimacy, in their understanding, is a normative notion whose relevance and applicability does not presuppose the existence (actual or hypothetical) of a public authority. This reading would then imply that legitimacy is a normative criterion that applies to institutional and non-institutional contexts alike. I want to offer a reason as to why even theorists committed to this view should not interpret the question of legitimacy as primarily concerning the justification offered to the anarchist about why there should be states in the first place. The reason is that the anarchical challenge they are considering and trying to respond to is, beyond the appearance of a genuine philosophical interest in the theme of anarchy, no other than an exercise of fashioning one's own favourite opponent.

What Nozick is doing in particular is creating his favourite, ideal anarchist and then engaging with the arguments the latter might advance. The ideal anarchist Nozick engages with is one that agrees with him in conceiving the state of nature in a Lockean fashion as including a framework of natural rights (derived from the 'law of nature') that are unenforced (Nozick, 2013 [1974], pp. 10-12). Even more controversially, he would agree with Nozick on interpreting natural rights as 'side constraints' (p. 29) on permissible action, rather than, for example, *pro tanto* considerations that can be outweighed by aggregative welfare considerations, and in believing that fair schemes of distribution do not give rise to duties of reciprocation (pp. 90-95).

But then, instead of showing the implausibility of anarchism, Nozick has instead convinced himself (or someone who subscribes to his own moral view) not to embrace anarchism. This is not a response to anarchism, as the anarchist might have all sorts of reasons to reject the institution of the state, beyond those based on natural rights that are captured in Nozick's argument. This type of argument is instead a version of the principle of political legitimacy, one that, it is worth noting, is not necessarily in contrast with the conditional interpretation ('if there is an authority requiring compliance...') I envisaged before.

A critic could here turn my argument against myself and advance the following reflection; isn't the anarchical challenge somehow presupposed by the very notion of political legitimacy? The conditional reading of the principle of legitimacy can indeed be seen to entail the opposite conclusions from the ones I presented before; in the same vein as a standard of legitimacy necessarily implies that there must be, at least hypothetically, a legitimate state, it also implies that there must be, at least hypothetically, one that is illegitimate. This is certainly undeniable, but allowing for

the possibility of having illegitimate states does not necessarily entail considering anarchism as a serious alternative.

Imagine I use a standard of legitimacy L_1 to assess the legitimacy of state s_1 and conclude that s_1 does not meet L_1 . What would be the normative implications of this assessment? If the standard L_1 is such that there is at least another state s_2 in the currently existing world that meets L_1 , or that I can imagine a state s_3 that meets L_1 and that, although currently inexistent, could be feasibly brought about in the foreseeable future, then the normative recommendation can only be that we should move s_1 to a condition that is as much as possible similar to s_2 or s_3 . In other words, if the standard of legitimacy adopted is such that it is already met by some states or can reasonably be met by others in the near future, there is no reason to give credit to anarchism as we already have an alternative, which is the passage from illegitimate to legitimate states.

If, on the other hand, we adopt a standard of legitimacy L_2 which cannot be feasibly met by any state in the near future, it seems that we have indeed reached one version of anarchism, the one that is usually called in the literature 'philosophical anarchism'.⁴ Certainly, this version of anarchism cannot entail by itself the belief that the state as an institution ought to be overthrown. The reason why the belief in the illegitimacy of all existing and feasible states does not immediately entail political anarchism is that the plausibility of political anarchism necessarily depends on a further premise, which is that life outside the state is inherently better than life in

⁴ (Wolff, 1998 [1970]) and (Simmons, 1979) are the canonical example. In (Simmons, 1987), A. John Simmons explains the anarchical consequences of his theory and distinguishes them from full-blooded political anarchism, which corresponds to the idea that human life can only flourish outside the authority of the state. Not only that, Simmons explains that his version of philosophical anarchism is consistent with the view that certain states can be justified and that this consequently 'gives us moral reasons to refrain from undermining it and will typically give us moral reason to positively support [it]' (Simmons, 1999, p. 753).

the state. In the absence of a positive evaluation of life outside the state, one is simply left with a belief that does not indicate whether the institution that is called illegitimate should be replaced by something else.

In the next chapters, I am going to accuse some standards of legitimacy prevalent in the political liberal literature of being of the L₂ kind, which cannot feasibly be met by any state. The conclusions in this section show, however, that, if L₂-like standards are problematic, this is not because of their having unpalatable implications in terms of the acceptance of the institution of the state. Anarchy, in its normatively interesting version, does not constitute a possible conclusion of the question of legitimacy and as such can be legitimately ignored in the discussion. Whatever conclusion we will reach about legitimacy, they will not compromise our belief that the modern institution of the state is indispensable for justice.

1.1.2 Two Senses of Political Legitimacy

What I have shown in the previous paragraphs is that in addressing political legitimacy, the separate question of the justification of the existence of the state can be legitimately neglected.⁵ The main conclusion I drew is that legitimacy cannot be equated to a condition in which having a state is better than no state at all. There might indeed be illegitimate states, i.e. states that do not meet a standard of legitimacy, that are still better than a condition of pure anarchy, and, one could add, there might even be some idyllic anarchical scenarios that are better than the most legitimate of states.

What could political legitimacy then come to mean, once it is decoupled from the question of the justification of the state? To answer this, we have to see firstly

⁵ For the terminological distinction between justification of the state and legitimacy, I owe an obvious debt to Simmons's treatment of the matter (Simmons, 1999). Simmons reconnects the distinction to Locke and notices how Nozick is oblivious to it.

what a state *is* and *does*. We can provisionally define the state as an institution that issues, *within a territory specifically defined*, commands that are presumed to be binding. The commands are presumed to be binding, it can be added, both for those that are defined as the state *citizens*, whatever this notion implies, and for those non-citizens that simply happen to stay or reside within the territory, although the differences in citizenship status will transmit to the specific obligations each is subject to.

To say that the commands the state issues are presumed to be binding equates to claiming that the state has *authority*. Joseph Raz famously defined authority as the fact that an institution is able to provide *pre-emptive reasons for action*: '*the fact that an authority requires performance of an action is a reason for its performance which is not to be added to all other relevant reasons when assessing what to do, but should exclude and take the place of some of them*'⁶ (Raz, 1986, p. 46). Authorities are therefore characterised by their providing immediate and content-independent reasons for action that those subject to it take as not merely adding to the reasons they already have in favour of performing an action, but *replacing* it. They follow the course of action dictated by the authority neither because they consider the course of action on balance the right thing to do nor because the fact that the act is recommended by an authority is a particularly weighty reason. Instead, they follow that course of action because, once a directive has been issued by an institution exercising authority, that is the *only* reason that they are bound to consider. In the opposite case, if those allegedly subject to the authority questioned every time

⁶I am here considering only one half of Raz's conception of authority, the one that we can consider as the 'descriptive' component. To this, Raz adds a normative component, the requirement that the pre-emptive reasons offered are reasons *dependent from*, although different from, some reasons for action those subject to the authority already possess (Raz, 1986, p. 47). This normative component turns authority into *legitimate* authority, thus providing the 'normal justification' for political authority I will discuss later.

whether a certain directive conforms to the reasons they already possess or adds something to them, there would be no authority at all.

Hence, what does it mean for an institution exercises authority over a certain population within a particular territory to be legitimate? It could only mean that it can exercise authority without interference or that it has a *right* to do so. We conclude therefore that a legitimate state is one that has a right to exercise authority over a defined population within a defined territory.⁷ I will now proceed to define a right of authority, starting from a taxonomy of the interferences that the right is supposed to limit.

Where could the interferences against the right to exercise authority come from? We can imagine two sources. A state can have *external interferences* on its right to exercise authority when entities outside the territory do not respect its possession of the right and instead reclaim that right for themselves. What is threatened through this kind of interference is the *exclusive* possession of the right to exercise authority of a certain state over a defined territory. We can call this kind of external interference violation of a state's *territorial sovereignty*. Correlatively, a state is not legitimate in this sense when it has no right to exercise exclusive authority over a defined territory, which is to say, it has no *justified territorial sovereignty*.⁸ This is the conception of political legitimacy that is usually in play in both international law

⁷ David Estlund has a similar view when he claims that 'by authority I will mean the moral power of an agent [...] to morally require or forbid actions from others through command. [...] By legitimacy I will mean the *moral permissibility* of the state's issuing and enforcing its commands owing to the process by which they were produced' (Estlund, 2008, p. 2. My emphasis). Similarly still, David Copp: 'When we evaluate a state for its legitimacy, our concern is to assess its moral authority to govern' (Copp, 2008, p. 4).

⁸ I speak of *justified* territorial sovereignty, instead of territorial sovereignty *simpliciter*, because I want to emphasise how my analysis has no empirical nature. If a state's territorial sovereignty is constantly violated, we can sensibly claim that the state has no *de facto* territorial sovereignty, but this would imply nothing about the legitimacy of its exercising authority over the defined territory. The opposite is also trivially true; a state can have *de facto* territorial sovereignty without having any right to exercise authority in that territory.

and global ethics debates,⁹ and, as I am going to show later, is the one that is (maybe surprisingly) less affected by the existence of disagreement.

The second source of interference is, on the other hand, *internal* and derives from the unwillingness of citizens and denizens of the ascribed territory to consider the reasons provided by the state as capable of conditioning their actions. What is here threatened is the possibility itself that the state can exercise authority. Correlatively, a state is illegitimate in this sense when it has no right to demand that citizens adjust their conduct in some sense following the directives it ordinarily issues. The state is *vice versa* legitimate in this sense when it can legitimately issue commands that are taken to be binding for those that reside in its territory, and whose reasons offered for compliance are, in the Razian sense, pre-emptive. This is the sense of legitimacy underlying the discussion of political legitimacy within classical liberal literature and today particularly relevant in the dispute between political liberalism and contemporary 'political realism'.¹⁰

The legitimacy of the state in this second sense is equivalent to its right to demand that those subject to its authority adjust their conduct following its directives, thereby showing that they consider the reasons offered by the state as pre-emptive for their action. This formulation however is insufficient, because it only refers to the right to authority in the abstract, without disclosing how the right can effectively protect the state from the internal interferences that threaten its exercise

⁹ For this use of political legitimacy, see especially (Rawls, 1999), (Cohen, 2006), (Buchanan, 1999) (Buchanan, 2011), (Meckled-Garcia, 2016).

¹⁰ For this second use, see especially (Raz, 1986, pp. 38-105), (Simmons, 1999), (Buchanan, 2002), (Estlund, 2008), (Quong, 2011) (Erman & Möller, 2015) for the 'liberal' view, and (Williams, 2005) (Rossi, 2012) (Horton, 2012) for the 'realist' alternative.

of authority. What is missing is the explication of an account of *political obligations* that have to be present so that the state can exercise its right of authority.¹¹

Not all authors, it must be noticed, concur that an account of political obligations is necessary to make sense of political legitimacy. Some of them dispute what William Edmundson (one of the critics) calls the ‘simple correlativity thesis’, holding that ‘a state is legitimate only if it claims to impose and in fact does impose on its subjects a general, at least *prima facie*, duty to obey its laws’ (Edmundson, 1998, p. 36). Allen Buchanan (Buchanan, 2002), another sceptic, even considers the question of political obligations as ultimately irrelevant in political theory. Hence, it is paramount that I first provide an explanation as to why political obligations and political legitimacy are, in my presentation, the two faces of the same coin.

Take Buchanan’s account holding that ‘an entity has political legitimacy if and only if it is morally justified in wielding political power’ (Buchanan, 2002, p. 689), or David Copp’s view that legitimate states are those that have a ‘sphere of privilege within which [they] could permissibly govern’ (Copp, 2008, p. 20). What these formulations seem to entail is that governing or, in Buchan’s words ‘wielding power’ – which he defines as the ‘monopoly, within a jurisdiction, in the making, application, and enforcement of laws’ – are, in normal circumstances, illegitimate acts that are therefore in need for justification whenever they are exercised. Interestingly, this move has blurred again the distinction drawn above between political legitimacy and justification of the institution of the state, as it assumes that the fundamental question of legitimacy in political theory is to show that the generally illegitimate ‘wielding of power’ is at least in some cases justified.

¹¹ My point here is purely contingent on the discussion. I do not want to venture into the discussion about whether *every* right necessarily implies correlative obligations, but just to show how the specific right to authority can only be meaningful if associated with a robust account of political obligations.

I have already expressed above my critiques to this type of reasoning. It assumes that the justification of the state has to be explained, in purely pre-institutional terms, against the background of a state of nature in which persons already possess natural rights, among which, presumably, is the right 'not to be wielded power' against.

If we subscribe to the *quasi*-Razian¹² interpretation of authority I presented above, a right against being subject to an authority is a contradiction in terms. Remember that authority here refers to the mere fact that an institution is capable of providing *what are taken* by the subjects as pre-emptive reasons for action. Now, how can I have a right against the fact that *I* take certain directives to be binding, in the sense that I believe the existence of the directive constitutes *for me* a pre-emptive reason to act in a certain way (or omitting to act in a certain way)? This would completely subvert the meaning of the concept of a right, as it posits a right *against* performing an action only the holder of the right herself can initiate. If what is meant here is that taking the directives issued by an institution as always capable of conditioning one's course of action is unacceptable, for instance because it undermines personal autonomy, then the language of duty (specifically, of duty to the self) would be more appropriate.

Why do these authors refrain from associating legitimacy and political obligations? The answer, I believe, resides in Edmundson's 'Strong Correlativity Thesis', which, an aspect I have so far neglected, is not exactly equivalent to the idea that *political* obligations are the direct correlatives of the right to exercise authority. The difference is that Edmundson seems to assume (before denying it later) that the best way to conceptualise political obligations and to make them normatively significant

¹² I say *quasi* Razian, because pre-emption is, for Raz, just one of the two fundamental features of authority. See note 6 above.

is by reducing them to a simple obligation of obedience.¹³ In that sense, legitimate states would be ones where citizens have an obligation to obey all the directives that the state issues. But, as Edmundson and many others notice, this seems highly implausible; so implausible that, were this the only way to define political legitimacy, philosophical anarchism (the idea that no existing or imaginable state is or will ever be legitimate) would become the default position.

Indeed, very few would happily accept that there is such a thing as an all-things-considered moral obligation to obey the law, with the reasons for this reluctance many and notorious. Firstly, very few would accept that citizens might have a moral obligation to obey the law (even of a *pro tanto* or *prima facie* kind) in circumstances where the state is gravely unjust. This is, however, the weakest objection as the advocate of Strong Correlativity could obviously reply that a gravely unjust state is then illegitimate. But, once we move away from Nazi Germany or Apartheid South Africa, further complications emerge. For example, it seems highly implausible that a state could be deemed illegitimate because of its enforcement of a single unjust law, particularly so if the law has been approved through perfectly legitimate procedures. Yet, it seems also implausible to claim that all citizens have a moral obligation to obey a gravely unjust law, purely in virtue of the fact that the authority requiring compliance is legitimate. Finally, even in case of just or justifiable laws, there is the problem of everyday compliance. It seems implausible that people are under a moral obligation to respect the law even in the case where one is certain a specific violation will not generate any negative consequences (not even the one of setting a bad precedent and rendering compliance with the law less common in the future).

¹³ Copp takes an altogether similar argumentative route: in rejecting the traditional correlativity view, he characterises as the view that the 'legitimacy of a state would consist in its subjects' having a moral *obligation to obey the law*' (p. 10 – my emphasis), instead of the view that the legitimacy of a state consists in its subject having an obligation to respect the authority of the state.

Suppose I find myself at a junction in the middle of the desert with no cars approaching (let's assume full visibility on all sides) and no witnesses that could take inspiration from my bad example. Am I still under a moral obligation to respect a 'stop' sign or a red light? Those who, like Edmundson, are puzzled by this example (Edmundson, 1998, p. 12) are probably confused by the presence of a justifiable law, within a legitimate state, that does not seem to provide decisive reasons for action.

An even more general problem is that the obligation to obey the law seems to imply a surrender of the moral agent's judgement. In case the state is legitimate, according to this unattractive model, the moral agent would be prevented from elaborating autonomous judgements about the appropriateness of following a norm. This seems to be an unwelcome, but inevitable, consequence, of explicating legitimacy at the domestic level as a right to exercise authority, at least as long as we remain committed to the Razian account with its definition of authority in terms of capacity to produce pre-emptive reasons for action. It might indeed seem that, if legitimacy implies political obligations, then citizens and denizens of a territory over which a state exercises authority legitimately have indeed to surrender their judgement on every matter that is regulated by a state's directive.¹⁴

This, however, is far from being necessary, as it can be shown through a minor modification of the Razian account. Instead of taking pre-emptive reasons to always be about the performance of a certain action (and, in case of prohibition, its omission), we take the pre-emptive reasons for action derived from an authority as counting in favour (and, through their pre-emptive character, replacing all other reasons possibly counting in favour or against) of a *disjunction* of actions. Specifi-

¹⁴ Raz rebuts the 'surrender of judgement' interpretation of authority in (1986, pp. 38-42), but my argument is only loosely related to his.

cally, the reasons offered by an authority recommend *either* complying with the directive *or* willingly accepting the consequences of non-compliance that the authority has previously established.

We therefore define this disjunctive duty to either comply or accept the consequences of non-compliance as the *content* of political obligations.¹⁵ We can immediately notice how this cannot by itself constitute a complete breach of personal autonomy, when we mean by this the priority of personal choice in the selection of one's favoured course of action. The individual who discharges political obligations is still free to choose whether compliance is the right course of action; the alternative of refusing to comply and accepting the consequences that this challenge of authority might generate is still consistent with respecting political obligations. The individual who willingly accepts the consequences of non-compliance is still letting her conduct be influenced by the directives issued by the authority, and as such is still taking the reasons offered by the authority to be pre-emptive for her action.

What about, then, those who refuse even to accept the consequences of non-compliance? First, it is crucial to qualify the second horn of the disjunctive duty, as my formulation could seem a bit vague. The most straightforward way to show that one is accepting the consequences of non-compliance is accepting to be punished. Civil disobedience is a classic case; the civil disobedient refuses to comply with a norm but still respects the authority of the state insofar as she allows her conduct to

¹⁵ The modification of Raz's account will allow me to show in the next paragraph that, *contra* Raz, civil disobedience and respect for authority are compatible. This could create a parallel with David Lefkowitz's argument (Lefkowitz, 2007) in favour of a right to civil disobedience (and I have to admit my debt to Lefkowitz for the term 'disjunctive' employed before). However, my argument merely entails that civil disobedience is a way to express respect for authority and that it is a moral right. It does not entail, unlike Lefkowitz's, that the state has no right to punish the civil disobedient; on the contrary, as I define the willingness to accept punishment as a condition for civil disobedience, the state must be in a condition to punish the non-compliant.

be conditioned by some directives issued by the state, specifically the ones about punishment for non-compliance.

It might here seem I have given a very partial representation of civil disobedience. Indeed, one might object, many civil disobedients only accept punishment for strategic reasons that have to do with the ultimate aim of their protest, for instance because the act of acquiescing to punishment publicly might have a powerful communicative effect.¹⁶ But they need not hold any belief about the overall legitimacy of the state system (that is to say, about the possibility that their conduct should be pre-emptively conditioned, in some regards, by norms of the state).

A possible response to this particular point is merely semantic. It consists in distinguishing between an act of civil disobedience, described as an act of refusal to comply with the law accompanied by the belief in the overall legitimacy of the authority of the state (which explains by itself the subsequent acquiescence to punishment), and an act of riot, characterised by the belief, on the part of the rioter, that the state has, on her, no authority whatsoever. This response is, I believe, theoretically sound, but the cost we would pay to commit to that, in terms of compromises to our ordinary language, might be too high. We do not want to refuse the status of civil disobedient to Rosa Parks, and call her a rioter instead, just because we are unsure whether she thought that 1955's Alabama exercised on her a legitimate form of authority. Furthermore, we want to keep using the distinction between riot and civil disobedience to elicit a difference in the character of the particular act that the protester undertakes, rather than in the belief she entertains. So that we can keep

¹⁶ David Lyons (Lyons, 1998) convincingly shows how neither Henry David Thoreau, nor Mahatma Gandhi, nor Martin Luther King, three of the most celebrated civil disobedients, ever thought that the authorities against which they were protesting deserved even a *prima facie* right of obedience. Their acceptance of punishment was merely, in their case, a communicative act.

saying, following ordinary language, that Rosa Parks was a civil disobedient, because, unlike, say, Fidel Castro in 1953's Cuba, she did not take weapons to demonstrate her opposition to the public authority, regardless of whether she thought the authority itself was overall legitimate.

But a better response is, I believe, available. I previously defined legitimacy, at the domestic level, as the right of a state to demand that its citizens or residents adjust their conduct following its directives, thereby showing that they consider the reasons offered by the state as pre-emptive for their action. I then described the duty to either obey the law or accept the consequences of non-compliance as the content of political obligations. A combination of these two sentences implies the following. A legitimate state is one that can legitimately demand that its citizens see the norms it issues as pre-emptively capable of conditioning their conduct. Citizens in a legitimate state are not required, however, to hold a further belief in the overall legitimacy of the state.

More importantly, however, I should emphasise that the model of legitimacy I am here presenting does not presuppose nor demand that legitimate states are ones where citizens hold a belief about the legitimacy of the state itself. Let alone that, the question of which states are legitimate is one I have not even started addressing. The question leading my investigation so far has been one slightly but significantly different, namely *what is* political legitimacy. My investigation so far has been mainly conceptual, in the sense that I have tried to defend a plausible interpretation of the *concept* of political legitimacy (thus providing what we can call a *conception*¹⁷ or account of political legitimacy). The question about which states are legitimate, although related, is different, as it concerns the particular property states

¹⁷ The distinction between concept and conception is famously elaborated by both Rawls and Dworkin in TOJ, p. 5 and (Dworkin, 2011, pp. 70-72) respectively, both of whom inspired by H.L.A. Hart (Hart, 1961, pp. 155-159).

need to exhibit in order to be called legitimate, or, in a terminology I now introduce, the *standard* of domestic legitimacy. The latter question is the one I will pursue in the next two chapters, as I believe that the theory of legitimacy through public justification is essentially a response to that particular problem.

We therefore conclude the discussion by recapitulating that political legitimacy is here understood to have two senses. In the first sense, it concerns the exclusive exercise of political authority, on the part of a single institution, within a certain territory. It is therefore equivalent to a right to *territorial sovereignty*. In the second sense, it is the right of an institution to issue directives that have to be taken by its subjects as capable of conditioning their conduct; correlatively, the right of an institution to be a legitimate source of *political obligations* is here taken to entail the disjunctive duty to either comply with the norms or willingly accept the consequences of non-compliance.

We further distinguished two ways to investigate the problem of political legitimacy. One question concerns the meaning of the concept of legitimacy, which, as we assume it is necessarily subject to diverging interpretations, can only be provided through a particular conception of political legitimacy. I define this as the problem of determining an adequate *account* of political legitimacy. A separate, although related, question concerns instead the concrete manifestations, so to speak, of political legitimacy, that is to say, the features or properties that legitimate states need to exhibit in order to be called legitimate. I define this as the problem of determining an adequate *standard* of political legitimacy.

Legitimacy

Domestic (right to issue binding directives)	Global (exclusivity in exercising authority within a specific territory)
Account (answers to the question about the meaning of legitimacy by providing a particular conception of the concept of legitimacy)	Standard (answers to the question about which states are legitimate by enunciating the properties that states need to exhibit in order to be called legitimate)

1.2 Legitimacy as Territorial Sovereignty and Disagreement

1.2.1 *On an Account of Legitimacy as Territorial Sovereignty*

Having presented a possible model of political legitimacy, it is now possible to reconnect this to the general theme informing the present work. The question I will now try to answer is whether and to what extent the fact of disagreement can constitute an obstacle to a state's achievement of political legitimacy. To do that, I will have to split the question following the two senses of political legitimacy I previously identified.

I will start with legitimacy understood as territorial sovereignty.¹⁸ A legitimate state in this sense is one that has a right to the exclusive exercise of authority within a specific territory; it is not subject to interferences from institutions that lie

¹⁸ Notice that my use of 'territorial sovereignty' is analogous to Rawls's notion of 'internal autonomy' throughout (Rawls, 1999).

outside that territory.¹⁹ The state might limit its sovereignty both externally, by voluntarily adhering to international treaties that partly transfer some of its competences to an international organisation to which it is a member (the euro-zone's monetary policy being a classic example) or internally, by devolving some powers to local authorities. These however are limitations of sovereignty that the state itself has adhered to (presumably in conditions that guarantee its freedom of choice), and as such do not represent a deprivation of sovereignty.

The interferences on sovereignty I am now considering are different as they are deliberate *violations* of the state's authority. The paradigmatic situation is that of an armed invasion, although this need not be the case, as we may consider as a violation of territorial sovereignty any attempt on the part of an entity that lies outside a state's borders to limit the exclusive right of the latter to exercise authority. What matters is that the violation of territorial sovereignty is accompanied by a recognition, on the part of the violator, that the state does not possess, or has lost, its right not to be interfered in the exercise of authority over its territory.

The question to be answered is therefore what grants the legitimacy of a state's exclusive authority over a territory. To answer this, we take the opposite perspective and consider instead what can justify a violation of a state's territorial sovereignty.

¹⁹ The numerous definitions of sovereignty in both international law and international relations (Philpott, 2016) are usually much more expansive than mine, as they would include, beyond the exclusive exercise of authority within a territory, also (1) the international recognition of the institution as a state ('statehood') and (2) the right to wield war. (1) will become relevant later, but at the moment it cannot be taken as a condition of sovereignty because we are concerned with what are the *normative conditions* that would make a state legitimate; the mere recognition of sovereignty would not do. (2) on the contrary can be left separate. If there is indeed a right to wield war, then territorial sovereignty might plausibly constitute a necessary condition for its exercise, but is certainly not sufficient.

The reason why I believe this perspective is preferable is that I am utterly sceptical about the possibility of justifying the right to territorial sovereignty in purely moral, pre-institutional terms, abstracting, that is, from the historical contingencies that led to the formation of the Westphalian system of nation-states. More specifically, I will now argue that theories of territorial rights that do not accept the conventional nature of territorial sovereignty face a dilemma. Either they are too abstract, and they turn into theories about the general legitimacy of states, as opposed to theories about the possession of territorial sovereignty, or they are historical and specific, which implies, however, that they risk deriving stringent normative implication from morally arbitrary contingencies.²⁰ Following a taxonomy that we can retrieve in A. John Simmons (Simmons, 2015), I will argue that both ‘functionalist’ and ‘voluntarist’ theories of territorial sovereignty (the latter of which Simmons himself subscribes to) are in fact indistinguishable from theories of political legitimacy at the domestic level, which means, in my account, theories about the justification of political obligations. ‘Nationalist’ theories, on the other hand, derive extremely stringent normative conclusions from often dubious empirical premises, and as such ought to be rejected.

Functionalist theories grant a state with the possession of territorial rights only to the extent that the state is capable of exercising a fundamental moral function, namely that of implementing justice and guaranteeing respect for individual

²⁰ An alternative to the dilemma would be to consider the territorial rights of states – or what I called a state’s right to territorial sovereignty – as merely ‘permissive’, e.g., as rights that allow a certain agent to commit an otherwise impermissible action, in view of the fact that they can help realise a state of affairs where the impermissible action occurs much less frequently. The proposal comes from Lea Ypi (Ypi, 2014); in her view, states’ exclusive authority over a defined patch of land can only be justified conditionally on the state’s agreeing to contribute ‘to the establishment of a global political authority realizing just reciprocal relations’ (p. 288). The proposal is interesting, but, *pace* Ypi’s own conclusion, does seem to lead to the conclusion that no state can today claim entitlement to any portion of the earth, even conditionally, until when it can prove it is taking steps to establish a ‘global political authority’.

rights. The capacity to implement justice, however, is a generic condition for the legitimacy (or, for Kant, the necessity)²¹ of states, and it does not vindicate by itself the special connection existing between a particular state and the territory over which it exercises authority.²² Functionalist theories, in their general form, do not show us what is morally wrong in annexing a foreign territory if the occupant state is a just one, nor do they tell us anything about the special link between a state and its people.

Functionalist theorists then have to include some additional features, other than the just exercise of authority, to explain that 'special connection'. Anna Stilz does that by including among her principles about the justification of territorial rights the fact that a state's 'citizens have a legitimate claim to occupy that territory' and that the state has not originated through usurpation of a previously existing authority (Stilz, 2011, p. 574). The problem is that, in order to avoid at all costs any dependence of the argument on historical contingency (and, as I show later, with good reason), Stilz ends up explicating the two criteria about specific territorial sovereignty in such vague terms that they can play only a minimal normative role. To exemplify, she writes that a person has a right to territorial occupancy if '(1) he resides there now or has previously done so; (2) legal residence within that territory is fundamental to the integrity of his structure of personal relationships, goals, and pursuits; and (3) his connection to that particular territory was formed through no fault of his own' (p. 585). This being the case, the right of occupancy is an attribute

²¹ 'A civil constitution, though its realisation is subjectively contingent, is still objectively necessary, that is, necessary as a duty' (Kant, 1996 [1797], p. 416).

²² I would argue that Kant's theory itself, as defended for instance in (Flikschuh, 2010), is a particularly weak basis for explaining the specificity of territorial sovereignty. The moral duty to leave the state of nature, which underlies Kant's justification for the existence of states (see Kant, 1996 [1797], p. 456), does not lead to any consequence regarding the particular allegiance that citizens owe to the state in which they happen to reside.

of every person who happens to reside in a certain territory due to no direct fault of hers and has developed there a web of connections that is central to her 'integrity'. These criteria are so normatively weak that, as Stiliz herself admits, they can only define as illegitimate situations of direct, and unjustifiable, military occupation. The descendants of the illegitimate invaders can however easily acquire occupancy rights as soon as they can prove that their residence there is 'fundamental to their integrity' and is not directly their fault. Hence, what Stiliz's criteria imply in most cases is that, if the state is legitimate (according to generic criteria of legitimacy), then it has justified territorial sovereignty over the territory it happens to possess, whereas, if it is not, it will by definition lack justified territorial sovereignty over *any* territory. In sum, the functionalist theory of territorial sovereignty is practically indistinguishable from a theory of state legitimacy at the domestic level.

The same objection applies to Simmons's own voluntarist model. In his case, the conflation between the two senses of legitimacy I previously distinguished is even more evident. Simmons is adamant in admitting that a state only acquires territorial rights through voluntary consent of its citizens (Simmons, 2015), exactly the same Lockean notion he had previously employed with regard to the general legitimacy of states (Simmons, 1999). He is thus led to the conclusion that virtually all existing states are then *both domestically and globally* illegitimate because neither their coercive authority nor their possession of a particular territory is the recognisable outcome of consensual allowances on the part of the citizens.

On the opposite end of the spectrum, authors belonging to the nationalist camp in contemporary liberalism advance a theory of sovereignty that is, at least in principle, better equipped to explain the special triple relation between a state, its citizenship, and its territory. In order to make their account compatible with the re-

ality of state formations, theorists who propose nationalist models of territorial sovereignty are not afraid of relying on historical contingencies. But historical contingency is doubly problematic as a ground for normative principles: it is often dubious, and, even when it is certain, it is by itself incapable of informing morally motivated reasons for action.

I will now clarify this trenchant remark. David Miller grounds his argument on a quasi-Lockean property view applied, however, to peoples instead of people.²³ In the same way as people acquire property rights, in the Lockean model, by mixing land with their 'labour', so peoples historically acquire territorial rights by 'improving' a territory they inhabit and making it 'valuable' (Miller, 2012, pp. 258-259). The first difficulty in an argument having this form is the problematic identification of what counts as a 'national people'. In particular, by arguing that a national people acquires sovereignty over a territory by editing the natural landscape and according value to some of its features, Miller seems to assume that all nations presuppose a pre-institutional people that had initially no entitlement over the territory but that then in the years following the foundation of the state acquired a right to it. But, if the national people is co-extensive with the entirety of the nation's citizenship, then what is even the point of using a controversial label such as 'people', as opposed to simply talking about the citizenship? If, on the contrary, the two are not co-extensive, the argument would lead to a further complexity, that is of specifying the exact subset of citizens who are entitled to territorial rights.

More importantly, however, the argument derives extremely stringent normative conclusions (namely, that no entity outside that state's territory can claim authority over any portions of the territory itself) from what is no other than a mere

²³ A similar appeal to a 'collectivist' interpretation of Locke's argument in defence of the right to property to justify a state's territorial rights is in (Nine, 2008).

historical contingency, which is that at some point in history a group of individuals started occupying a territory and subsequently began modifying its natural features and attributing symbolic value to it.

The first reason justifying cautiousness in the use of historical contingencies in political theory is that historical evidence is, often by its own nature, dubious and variously interpretable, and the various interpretations are very seldom neutral attempts at historical reconstruction. More often, historical re-interpretations will not be neutral at all and will aim instead at demonstrating, through a particular historical pedigree, the justification or lack thereof of a particular institutional structure.²⁴ At the very least, we can plausibly say that the risk of producing an historical reconstruction that is biased in favour of a substantive moral judgement concerning the current circumstances is always high.

This consideration, despite its being derived from empirical observation, cannot be dismissed too quickly. In fact, I cannot think of a single example of historical reconstruction, used by states to justify or celebrate their existence, that has not been hotly contested, very often with the contestations coming from marginalised groups that cannot buy the state's official narrative. The attitude of people in such marginalised groups will not be that of directly challenging the historical evidence but, which is even more detrimental for the achievement of legitimacy, that of refusing to recognise such evidence as a valid enough reason, or even a reason at all, to consider the state legitimate. People that end up asking, as Frederick Douglass famously did, 'what to the slave is the fourth of July?', do not question the veracity of a particular event that is taken to represent the genesis of the state, nor its ideal

²⁴ As Arah Abizadeh recognises, 'there is a tension between liberal democratic theory's commitment to norms of publicity, public justification, and freedom of expression, on the one hand, and the nationalist defence of a publicly shared identity dependent on historical myth, on the other' (Abizadeh, 2004, p. 291). Unlike Abizadeh, however, I do not take the main problem with historical justifications of territorial rights to consist in the fact that they are often based on fabrications of events. See later.

meaning. They believe, however, that, in view of the lack of concern they receive from the state, that original ideal is lost to them, and they fail to be moved by it.

Imagine, nonetheless, that we are dealing with a political community that universally recognises certain pivotal events in its history to jointly constitute the ground of the state's territorial sovereignty. Every single citizen recognises exactly the same historical evidence as both certain and capable of explaining why their compliance is due to the state in which they live, and to no other state. It remains the case that the main ground for the justification of the state's territorial sovereignty, which is accidentally accepted as valid by all those residing within the state's territory, is by itself a morally arbitrary contingency.

One could reply that the fact that citizens share that particular belief in the reliability on the one hand and motivational force on the other of the historical evidence is *not* an arbitrary factor and constitutes instead a morally salient fact. If that is the argumentative route followed, however, sovereignty is here ultimately justified by reference to a belief that is shared by the state's citizens (presumably in conditions that guarantee that their consent is sincere and unforced). In other words, what matters is that citizens believe in the validity of a justificatory strategy, broadly based on historical facts, adopted by the state.

The most plausible interpretation of the nationalist model has then given way to a resurgence of the voluntarist model *à la* Simmons. Despite its undeniable intuitive appeal, I keep seeing the voluntarist solution to the international legitimacy question as doubly problematic. On the one hand, as Simmons recognises, a purely voluntarist model can only lead, due to the empirical infeasibility of a universal consensus about the justification of the state's authority, to a defence of anarchism. Secondly, as I showed above, a voluntarist model fails to acknowledge the distinction

between the two senses of legitimacy I isolated before. This is not a merely conceptual problem. The inability at distinguishing the two senses of legitimacy implies the consequent inability at justifying the specificity of territorial sovereignty, that is to say, the reason why even just states can only claim authority over a particular portion of the earth's surface.

1.2.2 *On a Standard of Territorial Sovereignty*

So far, we have addressed the question of territorial sovereignty from a particular perspective, which is that of a plausible *account* of territorial sovereignty. Following the distinction at the end of the last section, I will now conduct the different investigation about a plausible *standard* of territorial sovereignty.

I will focus in particular on what are generally taken as the justifications for a violation of territorial sovereignty. The standard account in liberal theory holds that only gross violations of human rights (sometimes coupled with the violation of another state's sovereignty) can justify interferences of a state's territorial sovereignty. The linkage is so deep that, for authors subscribing to the so-called 'political' interpretation of human rights, it goes in both directions, as human rights are properly *defined* as those rights so central to human life that their violations can be a valid justification for interferences of a state's territorial sovereignty.²⁵

Among the defenders of the use of human rights as standards of territorial sovereignty, Joshua Cohen argues that what legitimises states is their valuing the

²⁵ 'I will take human rights to be rights which set limits to the sovereignty of states, in that their actual or anticipated violation is a (defeasible) reason for taking action against the violator in the international arena, even when—in cases not involving violation of either human rights or the commission of other offences—the action would not be permissible, or normatively available on the grounds that it would infringe the sovereignty of the state' (Raz, 2010, p. 387). Raz himself reconnects the 'political' view to (Rawls, 1999). For other explications of the view, which however do not share Raz's reductionism, see especially (Beitz, 2001).

membership of their subjects, in the sense that 'a person's good is to be taken into account by the political society's basic institutions: to be treated as a member is to have one's good given due consideration, both in the processes of arriving at authoritative collective decisions and in the content of those decisions' (Cohen, 2006, pp. 237-238). As long as states give due consideration to their citizens' welfare, which they can show by respecting basic human rights, they are protected from external interferences in their exercising authority.

Cohen does not neatly distinguish, however, between the two senses of legitimacy in the domestic and global sphere. Ronald Dworkin, by contrast, clearly relies on the distinction (although not literally) in his discussion of human rights. Dworkin's main concern is to distinguish human rights from the other 'political rights' that can also limit a state's implementation of its favoured policies.²⁶ Political rights in general are described as those entitlements respect of which demonstrates that a political authority treats its citizens with both equal concern and equal respect ('its policies treat their fates as equally important and respect their individual responsibilities for their own lives' – Dworkin 2011, p. 330). As the exact way to protect those political rights will presumably remain an object of disagreement for long, citizens of all countries have then a 'more fundamental, because more abstract, right [...] to be treated with the attitude that these debates [about political rights] presuppose and reflect – a right to be treated *as* a human being whose dignity fundamentally matters' (p. 335). A refusal, on the part of a state, to adopt an attitude of this sort while engaging with its subjects would gradually lead to that state's losing its right to territorial sovereignty. This would occur through actions, on the part of the international community, that vary from 'economic or [...] military interventions' in

²⁶ See the interpretation of 'rights as trumps' first defended in (Dworkin, 1997 [1970]). For the terminology of 'political rights' and their distinction from both legal and human rights, see (Dworkin, 2011, pp. 327-330).

the gravest cases, to the referral to 'international courts and tribunals', to 'more informal international pressure' in the mildest cases (p. 337).

The point had been previously argued, although with different terminology, by Rawls who claimed that human rights 'specify limits to a regime's internal autonomy' (Rawls, 1999, p. 79). Rawls cites among the 'urgent' human rights whose violation can compromise a regime's 'internal autonomy' those 'from slavery and serfdom, liberty (but not equal liberty) of conscience, and security of ethnic group from mass murder and genocide' (ibid.). Allen Buchanan also advances a similar view in his justice-based account of 'recognitional legitimacy'; he takes respect for basic human rights to be one of the components (together with the fact that a state has not come about through mere 'usurpation') of the 'minimal justice conditions' that can render a state internationally recognised as legitimate (Buchanan, 2011, p. 52). Interestingly, Buchanan explicitly cites Rawls's list of 'urgent' human rights to explicate the conditions of minimal justice.

The dependence of legitimate territorial sovereignty on respect for basic human rights raises two justificatory questions, namely why it is human rights in particular, as opposed to other moral notions, that play this normative role, and why the role is restricted to basic and 'urgent' human rights, which usually do not include for instance either most economic and social rights or the right to democratic institutions. What is, in other words, particularly significant about the rights that protect individuals 'from slavery and serfdom' to render their violation a legitimation for foreign interference, and why is not this role attributed to the entire set of human rights recognised by the Universal Declaration of Human Rights?

This question is directly relevant to the central theme of the present work. We can indeed observe that both levels of the basic human rights-dependence thesis, namely the appeal to human rights on the one hand and to the 'basic human

rights' in particular on the other, are, in conditions of disagreement, highly contentious. The contentiousness this question encounters is, if possible, even more intense than in the domestic sphere, as the discussion is now extended to the international arena, where we cannot even expect the sharing of a 'sense of justice' generated, in the Rawlsian picture, by the co-existence under the same, well-functioning, institutions.²⁷

The various responses provided in the literature are all, I believe, unsatisfactory, but at the same time helpful as they illuminate what is to me the most interesting feature of this sense of legitimacy: its unavoidably conventional nature. Following Charles Beitz (2001, pp. 272-273), we can distinguish two sets of justification that are used in the literature to support the idea that basic human rights can have this role in political theory. The first relates to their being, among the various notions present in a theory of justice, the one that is supposedly the least 'partisan' or 'parochial', to their being therefore accepted, although with variations, across cultural borders. We can call this alleged feature of basic human rights 'empirical popularity', and see it as the opposite of 'empirical contentiousness', that is, the mere fact that a certain value or judgement is considered controversial on a significant scale within a given environment.

The second set of reasons relates instead to the idea that basic human rights, regardless of their empirical popularity, *should* become the object of a consensus of reasonable agents in conditions of moral disagreement. This second type of argument is no other than an extension to the global arena of Rawls's constructivist methodology, with a different outcome derived presumably from the more intense kind of disagreement that we are bound to encounter at the international level. So,

²⁷ I will analyse the Rawlsian notion of 'sense of justice' in detail in § 5.3.

if at the domestic level conditions of reasonable disagreement would generate a consensus about principles of justice that, albeit limited to matters of 'basic justice', are still robust in their normative recommendations,²⁸ at the global level the consensus can only concern a minimal set of human rights. It is noteworthy how Rawls did not in fact take this route; instead, he explained the relationship between a country's 'internal autonomy' and its respect for basic human rights purely in terms of 'tolerance' of 'decent' regimes on the part of liberal societies.²⁹ However, it is a route that seems plausible and has indeed been advocated by some broadly Rawlsian authors.³⁰

I do not believe that any of the two arguments provide compelling reasons as to why respect for basic human rights can play the attributed role of guarantors of a state's territorial sovereignty. The first merely depends on the contingency that certain ideals are accepted as minimal standards of justice by peoples that would never reach an agreement on a more robust theory of justice. Because of the contingent unavailability of alternative standards, the different states must necessarily accept that the only standard of legitimacy to which they will hold each other accountable are those that pass this threshold of minimal agreement.

Of course, if this were indeed the case, holding each other accountable to minimal standards of legitimacy would still be better than not having any standards at all. Regardless of whether this is indeed the current scenario in the global arena, the

²⁸ The exact scope of the Rawlsian consensus on a reasonable conception of justice will be discussed in the next two chapters.

²⁹ The defence of this model is arguably one of the weakest examples of Rawls's 'reflective equilibrium' methodology: 'the reader has to judge whether a decent people, as given by the two criteria [one of which is respect for human rights], is to be tolerated and accepted as a member in good standing of the Society of Peoples. It is my conjecture that most reasonable citizens of a liberal society will find peoples who meet these two criteria acceptable as people in good standing. Not all reasonable peoples will, certainly, yet most will' (Rawls, 1999, p. 67).

³⁰ See for instance (Nussbaum, 2015).

recommendation to only hold each other accountable to whatever minimal standards of legitimacy it is possible to agree to could not possibly be part of a theory of international justice that has even minimal critical aspirations. This type of recommendation falls prey to the *status quo* bias objection, as it is incapable by itself of leading to a change in the moral standards adhered to by the various peoples. Imagine a significant number of members of the international community were unwilling to place an international ban on the practice of slavery; then a human right against slavery could never constitute a standard of legitimacy. Hence, in order to reach an internationally recognised standard of legitimacy (as we are assuming *any* standard is better than no standard), we would have to narrow down the scope of basic human rights until we reach a consensus. This would constitute at best a strategic response to the problem of disagreement in the international arena, but certainly not one morally informed.

The constructivist strategy is, on the contrary, undoubtedly a moral doctrine. I doubt, however, that it can provide a valid justification for the thesis that respect for basic human rights is a plausible standard of territorial sovereignty. The reason is simple; why would reasonable people adopt such a weak standard of legitimacy, instead of one that includes more robust principles of justice? We cannot merely assume that reasonable people place a normative relevance on the existence of national borders, and should never adopt a cosmopolitan perspective according to which the same standards of justice ought to apply across the globe. In order to exclude the cosmopolitan perspective, reasonable people would have to include significant empirical considerations in their deliberative reasoning; they would have to assume the necessity, or at least unavoidability, of national borders, and the impossibility of an international consensus on non-minimal principles of justice.

However, none of these empirical clauses is necessarily true and each of them could be false in imaginable circumstances. We can indeed well imagine, and with no excessive fantasy, both a borderless society and a world reaching a consensus on a theory of justice as robust, as, say, Rawls's justice as fairness. Notice that the principles of justice as fairness do not apply straightforwardly at the domestic level either; whatever we think of contemporary democratic societies, it would be hard to say that they have already reached the status of 'democratic well-ordered societies' of the Rawlsian ideal. Hence, if the lack of immediate application of an ideal principle of justice does not constitute a conceptual problem at the domestic level, it is unclear why it should be so at the international level. It follows that the asymmetry between principles of political legitimacy at the domestic and international level cannot be justified through a consensus of reasonable agents, as we can plausibly assume that reasonable people would adopt the same principles in both regards.

What may, on the other hand, justify the asymmetry is postulating the existence of two different forms of disagreements, one 'reasonable' that occurs at the domestic level within democratic well-ordered societies, and the other 'partly reasonable' that applies at the international level. Partly reasonable agents would be peoples or states that accept and share minimal principles of justice about the decent treatment of their subjects, such as Cohen's principle of 'membership'. These peoples or states would *in fact* treat their members with some equal respect, for instance by always presenting justifications for the coercive norms they enact, but they would not formally recognise equal respect nor provide a robust framework of rights that individuals can appeal to if they feel wronged by the authority. In such circumstances, one could indeed imagine an overlapping consensus of partly reasonable agents that would include the idea that a provision of basic human rights is necessary for political legitimacy.

But, again, the exact purpose of this shift of focus from reasonable to partly reasonable disagreement is unclear and could be interpreted differently. Specifically, either the less than fully reasonable character of disagreement at the international level is an empirical contingency (which we can, as such, empirically observe and measure), or it is an idealised abstraction loosely based on common sense. In the first case, the theory of overlapping consensus has acquired an empirical connotation and we are back at the ‘empirical popularity’ argument already discussed above. In the second case, on the other hand, I cannot see how the argument cannot avoid begging the question. To reiterate, we are questioning the use of basic human rights as a standard of international legitimacy (‘territorial sovereignty’). The response suggested by a reasonable consensus-kind model seems to be that partly reasonable peoples or states – defined as those that accept minimal principles of just treatment for their subjects – would tend to consider similar principles of minimal justice to hold each other accountable as standards of international legitimacy. In other words, the response about the standards of legitimacy would depend on the identification of the right kind of partly reasonable ideal person (or people) adopted. The features characterising part-reasonableness would be identical to the standards of legitimacy adopted as those will be, once again, the only standards people in disagreement hold each other accountable to. In other words still, we first define partly reasonable people as those that share no principles of justice other than basic human rights standard, and then build on this assumption a theory which unsurprisingly yields the result that basic human rights are the only standards of legitimacy at the international level that partly reasonable people hold each other accountable to; quite the paradigm of an *ad hoc* construction!³¹

³¹ Charles Beitz, who defends Rawls’s conception of human rights, clearly recognises this problem. ‘It is part of the definition of decent (and liberal) societies that their institutions respect human rights; although there is in this respect an agreement about human rights, an appeal to this agreement to

I propose starting from a different model. I suggest starting from the acknowledgement that the existence of the Westphalian nation-states that claim territorial sovereignty over a determined patch of land is a historical contingency that does not generate, by itself, restrictions in the formulation and justification of principles of justice *and* legitimacy. That is, at the ideal level, the same principles of justice and legitimacy apply both at the domestic and global level. These ideal principles of justice and legitimacy have however a primarily aspirational character; as such, they do not delineate what justice requires in all possible circumstances.³² Now, in order to see what justice requires in ordinary circumstances, one would need to take into account some familiar facts about the world that were set aside at the level of abstract elaboration.

My suggestion is to consider the existence of the nation-states reclaiming territorial sovereignty among these 'familiar facts'. Notice that I am not taking the existence of political institutions *per se* among the facts that an ideal theory of justice can legitimately ignore, because, in doing that, I would fall prey to the anarchical fallacy I presented at the outset of the chapter. I am simply considering the currently existing system of nation-states and international relations for what it is, an historical contingency. It follows that the normative instruments that have been progressively adopted to regulate that system are also an historical product, and specifically the product of a particular history of international relations.

The use of a certain subgroup of human rights as the standard that guarantees respect for a state's territorial sovereignty is a policy that has a specific historical pedigree, which we can reconstruct following the history of international relations and the formation of international organisations in the last century. Without

explain the authority of human rights or to determine their proper scope would be circular' (Beitz, 2009, p. 98).

³² See my defence of 'ideal theory' before (§0.3).

that specific historical pedigree, both justificatory questions I posed before (why human rights, and why basic human rights specifically) remain unanswered. If we try, on the other hand, to answer those questions at the level of purely abstract normativity, wondering why it should be the case that, in any imaginable circumstances, basic human rights could have that normative role, we commit a category mistake. That is because there is simply no abstract, fact-insensitive reason why that should be the case.³³

1.2.3 Conventions and Reasonable Disagreement

We can define normative principles that are mainly grounded in particular social facts (such as the presence of a shared social practice or the previous stipulation of an agreement) as *conventional* principles.³⁴ Territorial sovereignty is, in itself, a conventional normative notion, because it depends on the convention, which gradually became accepted in international relations, that the exclusive authority of a single state over a single territory, regardless of how that originated, is inviolable, unless the state has committed a significant wrong either in the relations with its neighbours or in the treatment of its subjects (which is where the basic human rights thesis specifically applies).

³³ Notice that my point here purely concerns the use of basic human rights as a standard of territorial sovereignty, which, I argue, is no other than the outcome of a particular practice in international relations. I am not trying, on the other hand, to *define* the *concept* of human rights starting from that practice, as authors such as Raz (see note 27 above) or Beitz (Beitz, 2009) try to do.

³⁴ For the distinction between moral and conventional principles, I am greatly indebted to Nicholas Southwood. In (Southwood, 2011), he indeed suggests, against models based, respectively, on either the form or the content of moral judgements, that what distinguishes conventional principles is the role played by conventions in their grounds. I have however translated Southwood's reflections from normative judgements to normative principles.

Notice that the authority of a state over a territory is not questioned, unless the state itself has committed some prominent wrongs, even when the state's possession of at least some part of the territory is notoriously due to what would today be considered an injustice. Most would consider today the colonial way of appropriating territories as generally unjust, barred maybe those territories that were completely uninhabited. However, if you take a state such as the US, almost the entirety of its current territory is in some way or another the result of colonial modes of appropriations, either directly (the lands that were gradually conquered by fighting the natives in the 19th century) or indirectly (the territories, such as Florida and Louisiana, that were acquired from previously existing colonial empires). Nonetheless, the recognition of such injustices in the original appropriation of the territory over which the state currently exercises authority does not, by itself, put into question the inviolability of sovereignty.

By underlining the primarily conventional nature of territorial sovereignty, I am not denying that territorial sovereignty might be, in more than one sense, morally valuable. There are of course numerous advantages that derive from the international recognition of territorial sovereignty, from the prevention of a significant number of conflicts to the limitation of powerful nations' imperialistic ambitions. These advantages provide, however, only partial justification for the normative cogency of territorial sovereignty, for they could have been obtained, following a different configuration of events, without its recognition. For example, in an entirely borderless society, conflicts over the possession of territory would be similarly avoided and there would be no powerful nations that would attempt dominion over the others. As already argued previously, therefore, what ultimately justifies both the normative cogency of territorial sovereignty and its dependence on respect for

basic human rights is the fact that the view has attracted a stable consensus, among sovereign states, roughly after the end of the Second World War.

What is the bearing of these reflections on the question of the relationship between legitimacy and disagreement? Simply, because of the conventional nature of the respect-for-human-rights standard that is currently employed to ground legitimacy at the global level, the justification of the standard itself cannot be subject to reasonable disagreement. To see that, consider the burdens of judgement, which, to reiterate, are defined as those impediments in the ordinary use of the faculty of judgement that prevent a condition of agreement from coming about in certain areas of reflection (roughly, in the realm of values). The role of the burdens of judgement is limited here by the fact that the justification of territorial sovereignty depends on the existence of a convention, which is an easily recognisable fact and not one where the 'evidence is conflicting and complex' (PL, p. 56).

We can expect reasonable agents to disagree about the morality of the convention; we can further expect reasonable disagreement concerning its best interpretation of the convention itself. Both discussions will probably lead to important conclusions regarding whether the convention ought to be substantially changed, or even rejected altogether, or whether it can still be maintained as it currently stands.³⁵ As long as the convention is still part of the consensus, however, there can be no reasonable disagreement about the fact that the convention is, *at least in one sense*, still normatively valid.

³⁵ Using Dworkin's interpretivism about social practices as a paradigm, reasonable disagreement about a particular conventional norm can encompass both the 'interpretive stage' where the 'interpreter settles on some general justification for the main elements of the practice' and the 'post-interpretive or reforming stage, at which he adjusts his sense of what the practice "really" requires so as better to serve the justification he accepts at the interpretive stage' (Dworkin, [1986] 1998, p. 66).

To understand that ‘sense’, we have to embrace a minimal form of legal positivism,³⁶ namely one holding that the *conventional* validity of a conventional principle purely depends on the existence of a social fact, which in our case corresponds to the international consensus about both the value of territorial sovereignty and its base in respect for human rights. Rejecting this minimal form of positivism would amount to the view that there can be no sense of validity that includes purely conventional considerations; validity, according to the objector, would *always* incorporate non-conventional considerations, such as whether the norm is consistent with morality or rationality or common sense. The objection is conceptually sensible, but can be rebutted through a purely semantic move; as my point is conceptual and not simply grounded in linguistic analysis, what I called ‘conventional validity’ can be renamed ‘conformity with the consensus’. It would remain the case, however, that there is an important sense of evaluating conventional principles that transcends from any non-conventional feature, such as conformity with morality or rationality.

These reflections bring me to the final point in this chapter, which is that, in light of the conventional grounds of the principle of territorial sovereignty (holding that states can retain the exclusivity in authority over their territory insofar as they

³⁶ H.L.A. Hart described the positivism of Jeremy Bentham and John Austin, which he spent great part of his career re-elaborating, as consisting of two theses, ‘first, in the absence of an expressed constitutional or legal provision, it could not follow from the mere fact that a rule violated standards of morality that it was not a rule of law; and, conversely, it could not follow from the mere fact that a rule was morally desirable that it was a rule of law’ (Hart, 1958, p. 599). The positivism I am interested in here, however, is concerned with both the existence and the validity conditions of conventions, rather than mere laws.

A version of positivism that could support my thesis here is, for example, the ‘inclusive positivism’ that Kenneth Eimar Himma defends, with its Weak Conventionality Thesis: ‘What explains the authority of the validity criteria in any conceptually possible legal system [...] is that such criteria constitute the terms of a social convention among the persons who function as officials’ (p. 129). Inclusive positivism includes for Himma the idea that ‘there are conceptually possible legal systems in which the validity criteria include substantive moral norms’ (p. 136); my point here is different, insofar as I am distinguishing *two senses of validity* (conventional and non-conventional) that may apply, ideally, to all conceptually possible conventional normative systems.

do not violate basic human rights and do not assault other states' sovereignty), reasonable agents may reasonably disagree only on whether the convention should have occurred in the first place, but not on whether it is still currently existing *as a convention*.

It follows that the notion of public justification, which in political liberalism serves the purpose of reconciling legitimacy and reasonable disagreement, cannot have a significant role on the global side of legitimacy. The justification that one can offer for the dependence of territorial sovereignty on respect for basic rights is necessarily rooted in some social facts; facts that can be criticised from a moral viewpoint, but cannot be rejected as irrelevant when they constitute the reason why we have a convention in the first place. A fact-rooted justification cannot, therefore, address reasons that derive from citizens' comprehensive doctrines, which is precisely what, on the other hand, a public form of justification is supposed to do, as I am going to display in detail in the following two chapters. Hence, when political liberals take public justification as the ideal solution to the question of political legitimacy, their focus must be exclusively on the domestic side of the question.

2. The Failure of Public Justification

The last chapter concluded with a crucial point regarding the relationship between legitimacy and public justification: due to its conventional nature, legitimacy at the global level is not affected, at least in a philosophically interesting sense, by the existence of reasonable disagreement. When Rawls claims, in the passage I have now extensively quoted, that a state is legitimate only when its 'constitutional essentials' are expected to be 'reasonably acceptable' to free and equal citizens, he must be using legitimacy in a sense other than territorial sovereignty. Similarly, when Nagel defines the issue of 'political legitimacy' as 'the history of attempts to discover a way of justifying coercively imposed political and social institutions to the people who have to live under them' (Nagel, 1987, p. 218), he cannot mean by 'legitimacy' anything even remotely related to international relations.

If that is the case, the only possibility is that public justification is to be interpreted as a solution to the question of *domestic* legitimacy. Some political liberals, including Rawls and Nagel, sometimes seem to use the principle of legitimacy through public justification in a kind of axiomatic way, as if the belief in the necessity of public justification for the satisfaction of political legitimacy was simply one of the unanalysable principles of a liberal theory of justice. The question of whether a theory of justice might legitimately include axioms is a methodologically complex one that I do not intend to address here. However, I believe I gave sufficient reasons in the last chapter to explain why the principle of legitimacy through public justification cannot be treated as an axiom, and why, if we want to be charitable to its advocates, we cannot interpret it that way. The reasoning is that, in order to make sense of the concept of political legitimacy, we must offer a conception of it, or, in the terminology I previously employed, an *account*. That account will then identify the moral conundrum political legitimacy is supposed to address.

One moral problem we come across in discussing the state is coercion, and indeed some scholars, those following Gerald Gaus and his ‘justificatory liberalism’ paradigm, seem to treat the question of the state’s legitimacy as if it primarily concerned the justification of the coercion exercised by the state. But that cannot be right, and, again, reasons of exegetical charity require that we interpret even the moral concern underlying ‘justificatory liberalism’ differently.¹ As Gaus himself recognises, the state is not the only entity that can legitimately exercise coercion; in fact, morality itself has a coercive (or, as Gaus writes, ‘authoritative’) character, insofar as it requires to dictate human conduct.² Hence, the two questions of legitimacy and justified coercion necessarily diverge, although their solution might be similar.

The identical treatment of the two separate questions of justified coercion and political legitimacy is wrong, however, for another reason, somehow symmetrical to the one just presented, which is that the state is not *just* a coercive entity, and therefore a legitimate state is not just a legitimately coercive entity. The previous chapter explained why. The state is not an exclusively coercive entity because the authority of a state, especially so when the state is legitimate, does not lie in the prospect of coercing those that disobey the state’s directions, but in the fact that the state provides reasons for action that those that fall under its jurisdiction take as pre-emptive for their conduct.³

¹ As a student of Gaus himself, Chad Van Schoelandt, does indeed in (Van Schoelandt, 2015).

² This seems to me the central theme of the magisterial (Gaus, 2011). See also how Gaus defends his account by relying on the idea that his ‘convergent normativity’ can solve the problem of the authoritative character of ‘social morality’ in (Gaus, 2015).

³ This point about the coercion of the state applies, if possible *a fortiori*, to morality at large. Even though morality does pose demands on moral agents, it does that by providing reasons for action, and not through an enforcement mechanism that might include threats for non-compliers. See David Enoch’s reply to Gaus in (Enoch, 2013).

That is why, in the previous chapter, I concluded my preliminary discussion by arguing that legitimacy at the domestic level amounts to a right to issue political obligations. Legitimate states are not those that exercise coercion legitimately, but those in which political obligations exist, and are *morally* binding for those that are under the state's jurisdiction. The question of moral bindingness is of the utmost importance here; political obligations, unlike legal obligations, are, in my account, a subset of moral obligations, that is to say, of the obligations we derive from and justify through moral reasons. Now, the burden on those defending the 'legitimacy through public justification' model is to demonstrate that these conditions (existence of morally binding political obligations) occur in states where the public justification requirement is satisfied. The entirety of this chapter is dedicated to showing why this is not the case.

I will first lay out what I take to be the principles characterising the 'public justification' paradigm in liberal political theory. I will dedicate the first section of the chapter (2.1) to an analysis of the different models of 'convergence' and 'consensus' in public justification, and take issue in particular (2.2) with the first. Against public justification understood as the non-dissent of all those members of the citizenry who can express 'intelligible' objections against public norms, I will raise both epistemological and moral doubts. After having introduced some points about the extreme epistemic demandingness of the position, I will accuse their defenders of cultivating a purely fetishistic interest in the theme of public justification.

I will then dedicate the last section (2.3) to the 'consensus' model of public justification, again relying on both epistemological and moral concerns. My aim is disputing the role that an idealised consensus of reasonable viewpoints might play in the achievement of political legitimacy. I will notice how a consensus of reasona-

ble doctrines tends to generate its own form of 'interpretive' disagreement. A reasonable consensus can only work as a solution to the problem of legitimacy if the normative implications of interpretive disagreement are shown to be substantially different from those of non-interpretive, 'substantive' disagreement. But, as I will show, once we analyse how interpretive disagreement plays out in the context of everyday norms, we see that its outcome is as detrimental for political legitimacy as that of substantive disagreement.

The discussion will continue in the next chapter, where I will devote my analysis to the notion, which I have used so far uncritically, of reasonableness. The conclusions I will raise there are going to be equally sceptical. I will argue that the notion of reasonableness is morally vacuous and, if taken seriously on an epistemological perspective, leads to a form of scepticism, which is a paradoxical result for political liberalism.

The conclusions of this and the next chapter will be globally negative; I will not accompany them with a positive thesis about political legitimacy. The reason for this restraint lies in the general theme of the present work. Legitimacy only interested me insofar as it proved to be relevant to the normative treatment of disagreement in liberal societies. Once I have shown this is not the case, I can move on to other normative notions that I will show are more relevant to the question. The question of legitimacy will be then left open, whereas one particular answer to the question, the one centred on the notion of 'public justification' (or 'public reason') will be dismissed.

2.1 Two Models of Public Justification

'Public justification' can be first described as a model that public institutions can rely on in their attempt to legitimise their authority. So described, it aligns perfectly with

the long project of liberalism, the one Waldron identified as the demand that the 'social order should in principle be capable of explaining itself at the tribunal of each person's understanding' (Waldron, 1987, p. 149). But the peculiarity of the 'public justification' paradigm lies in the type of constraints it introduces for a justification offered by the state to be 'public'. The introduction of these constraints marks an evolution in the thought of John Rawls, who became disillusioned the justification he himself had offered regarding his theory of justice as fairness⁴ could attain that state of publicity necessary to legitimise states in the real world.

Hence, the introduction of the 'shared-reasons' constraint on public justification, holding that a model of justification is public, and therefore capable of legitimising the authority of the state, only if based on reasons that are shared by, or easily accessible to, the political community at large. For Rawls and his most faithful followers, the reasons offered by the state, to be public, need to seem acceptable to a particular kind of citizen, which we can variously interpret as an idealisation of the liberal citizen or a subset of the actual denizens of contemporary democratic societies: the reasonable.

The selection of each interpretation will give rise to its own set of problems. In case the reasonable are interpreted as a kind of ideal model of liberal citizens, a critical question immediately springing up is whether this idealisation is, in the context of public justification, appropriate.⁵ If, on the other hand, one would take the reasonable to be a particular subset of the actual citizens of contemporary democratic societies (on a par, for example, with such categories as 'the religious' or 'the

⁴ In *A Theory of Justice*, Rawls's main preoccupation, once he has defended justice as fairness as a model on which philosophers can reach 'reflective equilibrium' (TOJ, 18), is that of showing that justice as fairness is 'congruent' with a particular theory of the good of a broadly Aristotelian outlook (I will work on this in Chapter 5).

⁵ For a view that conditions the legitimacy of idealisations on the context of the theoretical problem one is trying to solve, see (Enoch, 2015, p. 118 ff.).

disabled'), then the problem would be that of explaining why the possession of reasonableness plays that crucial normative role (namely, it entitles people to a scheme of public justification that is accessible to them).

We can leave these problems provisionally aside, as I will dedicate the entire next chapter to a critical analysis of reasonableness, and move instead to a different question. In what sense, we might wonder, are reasons that the state uses to justify its authority *accessible* to reasonable citizens? The model Rawls and his followers advance is that of an 'overlapping consensus' of reasonable doctrines. In Rawls's own theory, the 'overlapping consensus' plays the role of the ultimate mechanism for justification, which ensures that a conception of justice, such as justice as fairness, the one he proposes, can acquire 'full justification' for each citizen and, following from that, 'public justification' in the entire polity.⁶ As in the present investigation we are not interested in the justification of a particular conception of justice, but in the *use* of such conception in a state's attempt at legitimising its authority, my presentation of the model will follow this different aim. The reasons that the state offers are accessible to citizens not merely because they do not include reference to controversial conceptions of the good that some citizens might legitimately refuse to endorse (*pro tanto* justification),⁷ but also because they are compatible with the

⁶In the 'Reply to Habermas', Rawls defines 'full justification' as the moment where a political conception is 'embedded' or 'inserted as a module' into a particular citizen's set of commitments, whereas 'public justification happens when all the reasonable members of political society carry out a justification of the shared political conception by embedding it in their several reasonable comprehensive views' (PL, pp. 386-387).

⁷'In public reason the justification of the political conception takes into account only political values, and I assume that a political conception properly laid out is complete. [...] But since political justification is *pro tanto*, it may be overridden by citizens' comprehensive doctrines once all values are tallied up' (PL, 386).

reasons that feature in those very same reasonable conceptions (positive or full justification).⁸ Reasonable conceptions of the good, in sum, overlap on certain justificatory reasons that their holders take as sufficient to legitimise the authority of the state. As such reasons are part of the comprehensive doctrines that citizens have willingly adopted, and these conceptions of the good in turn provide the fundamental values that inform their entire life, citizens take these reasons to be not merely *pro tanto* valid, but decisive.

The overlapping consensus mechanism proves that citizens can retain their allegiance to the public authority not because they are coerced nor for prudential reasons of *modus vivendi*, but because they have *valuable* reasons,⁹ derived from their own value-providing conceptions of the good, that direct them to behave that way.

The model of ‘overlapping consensus’ so far presented has come under attack for a plurality of reasons.¹⁰ The critiques are kindled by the infinite ambiguities of the Rawlsian text. One of these ambiguities is the conflict between two apparently opposite normative requirements. On the one hand, when we consider the *pro tanto* or ‘political’ justification of a conception of justice (or, in my investigation, the use of that conception in the legitimation of a state’s authority), that conception must be justified in a way that abstracts from controversial values that some people might

⁸ For this political-legitimacy-driven reconstruction of the ‘overlapping consensus’, I am following Rawls himself (see above), but also Jonathan Quong’s presentation of the ‘common view’ of the overlapping consensus (later criticised) in (Quong, 2011, pp. 163-166).

⁹ I use ‘valuable’ reasons, instead of the frequently used ‘moral reasons’ (see, e.g., Quong, *passim*, or Scheffler 1994), because what specifically characterises the model of overlapping consensus is that citizens end up complying with the rules of the political authority for reasons that derive from their conceptions of the good, which in turn act as the vast containers of value (moral, but presumably also aesthetic) for their lives.

¹⁰ A critique I will not focus on, which proposes to change the role of the overlapping consensus from the ultimate justification mechanism to a strategy for identifying initially the constituency of liberal people to which justification will be made available, is in (Quong, 2011, pp. 180-191).

reasonably reject ('doctrinal abstinence'). When we move to the 'full' (for each citizen) or later 'public' (for all citizens) justification, however, the constraint is that the principles composing the conception of justice must be shown to be compatible with a plurality of reasonable conceptions of the good ('doctrinal compatibility'). To be fair to Rawls, the two constraints of 'doctrinal abstinence' and 'doctrinal compatibility' are not only logically compatible but, in his intentions, mutually supportive. Precisely because a conception of justice is not justified relying on controversial values, Rawls would say, we can show it is possibly congruent with a plurality of conceptions of the good.

But the *practical* compatibility of the two constraints creates some problems. If citizens see the conception of justice as a 'module' that can fit into their wider conception of the good, then presumably they will embrace that conception for non-political reasons, presumably because it fits well with the other (controversial) values they endorse. However, there is at least one sense in which citizens do need to understand a conception of justice in a 'political' way, which means, in the technical sense of 'political' Rawls adopts, in a way that abstracts from controversial values. Rawls argues that 'since the exercise of political power itself must be legitimate, the ideal of citizenship imposes a moral [...] duty – the duty of civility – to be able to explain to one another on those fundamental questions how the principles and policies they advocate and vote for can be supported by the political values of public reason' (Rawls, 2005 [1993], p. 217). The 'duty of civility' therefore mandates that citizens abstain from relying on possibly controversial values, derived from their conceptions of the good, when they engage in public discourse at least regarding 'fundamental questions'.

This creates some puzzlement; on the one hand, citizens can accept the political conception of justice for reasons that have nothing to do with its 'political' (i.e.

controversial-value-avoidant) character, on the other, they need to be prepared to abstain from those controversial values themselves when they engage in public discussions. Again to be fair to Rawls, I am not here accusing his presentation of the reasonable to have a schizophrenic or hypocritical character. In fact, concealing to others the real reasons why we have adopted a certain principle, when we believe those reasons are unappealing to them, is a common component of political strategy, or, even more widely, of the art of persuasion. But, if I want to be reasonable in the Rawlsian sense, I do not merely avoid mentioning controversial values publicly just because I know others are not buying them (that would be a merely strategic point), but because a duty of justice (the 'duty of civility') demands it so.

Furthermore, I cannot simply adapt my presentation of the conception of justice to the different audiences I am addressing (as any political strategist would advise me to do), but I need instead to rely always on the same set of considerations, namely on those considerations that reasonable people will accept. These reasons compose, unsurprisingly, a somewhat limited batch. The considerations that we can derive from the reasonable consensus are those that reasonable people are generally supposed to embrace, that is to say, respect for others' freedom and equality and the acknowledgement of the relevance of reasonable pluralism.

I now want to focus on an alternative presentation of political liberalism. Proponents of so-called 'justificatory liberalism' or the 'convergence' interpretation of political liberalism are sceptical about whether showing that the reasons offered by the state to justify its legitimacy are congruent with reasonable conceptions of the good is sufficient to show that the model of political legitimacy takes the fact of pluralism seriously enough. They suggest instead a model in which public justification only obtains when (mildly idealised) citizens can advance no 'intelligible' reasons,

derived from their own evaluative standards, to reject a particular norm the state has enacted.

'Justificatory liberalism' does away with three important requirements of public justification characterising the Rawlsian conception. Firstly, it does away with what Rawls defined in the 'Reply to Habermas' as the *pro tanto* type of justification, demanding that a conception of justice is justified only through reasons that reasonable people must accept. It then does away with the requirement that, when they engage in public reason, citizens can only refer to that set of reasonably acceptable reasons and not to their own controversial values, regardless of whether those same controversial values were instrumental in their adopting the conception in the first place. Finally, it practically (although not always literally)¹¹ does away with the distinction between reasonable and unreasonable citizens altogether. The ideal citizens that advocates of 'justificatory liberalism' take as the recipients of public justification are what Gerald Gaus calls Members of the Public, for which he means 'rationalized counterparts of real moral agents' (Gaus, 2011 b, p. 267). Members of the Public are actual agents taken with their real values and commitments, whose views have been tidied up to get rid of obvious inconsistencies and stretched carefully to draw the necessary implications. To take an example, if actual agent Jane believes that murder corresponds to the taking of a human being's life and that foetuses are human beings like any other, then Member of the Public Jane* will also believe (regardless of whether real-life Jane has consciously made that connection) that abortion is equivalent to murder.

¹¹ Paul Billingham (Billingham 2017, 545), for example, refers to the generic moral requirements for participating the public debate within the 'justificatory' versions of liberalism as requirements of 'reasonableness', even though he then moves to explain how they differ from the Rawlsian model. See also (Vallier, 2016 a, pp. 609-611).

What implications derive from these three significant departures from the Rawlsian model of public justification? Two particularly stand out. The first is that public justification of the authority of the state, which we are still treating as a prerequisite for the state's possession of legitimacy, becomes comparably harder to obtain. If we assume that the ideal recipient of a scheme of public justification is not any longer the reasonable citizen, with her commitment to robust, albeit basic, liberal values, but a Gaussian 'Member of the Public' with her actual commitments, then the scenario where one or more 'Member of the Public' will have reasons to reject that scheme become much likelier. At the same time, however, citizens who intend to engage in public reason, not subject any longer to the Rawlsian 'duty of civility', are not required to refrain from publicly using the values that derive from their own controversial affiliations. In a slogan, we could then say that 'justificatory liberalism' makes it harder for the state and easier for the citizen when it comes to public justification.

In order to prevent these implications from escalating to a condition of 'justificatory impasse', where public justification becomes impossible, advocates of 'justificatory liberalism' still suggest that the reasons that Members of the Public can advance to contest the justification of a public norm must meet certain criteria. Kevin Vallier proposes that public reasons must be 'intelligible', arguing that 'A's reason R_A is intelligible to members of the public P if and only if P regards A as entitled to affirm R_A according to A's evaluative standards' (Vallier, 2016 a, p. 603).

The requirement of 'intelligibility', which is a revised version of Gaus's original requirement of 'sincerity',¹² does not demand from citizens that they commit to basic liberal values. It does not assume that a liberal consensus is necessary in order

¹² 'Betty's argument justifying N to Alf is sincere if and only if (1) she is justified in accepting N; (2) she has a justified belief that N is justifiable in Alf's system of reasons and beliefs' (Gaus, 1996, p. 141).

to create a constituency of possible recipients of public justification, as in Jonathan Quong's reformulation of the 'overlapping consensus' argument (Quong, 2011, pp. 180-191). Nor does it assume, as in the more ambiguous Rawlsian formulation, that the constituency that is entitled to receive an 'accessible' scheme of public justification must accept certain restrictions on the type of arguments they can publicly advance. The only restriction 'intelligibility' imposes is that others must be capable of understanding the relevance of the advanced consideration for the evaluative standards of the citizen who advances them.

2.2 A Critique of Convergence

2.2.1 On the Epistemological Demandingness of Convergence

'Intelligibility' might well be the most refined version of the public accessibility requirement in 'justificatory liberalism', but introduces two prominent problems.

The first problem is epistemological. Vallier clarifies that his 'intelligibility' requirement does not introduce any agent-insensitive or 'externalist' considerations (Vallier, 2016 a, pp. 605-606). When I judge whether my co-member of the public Jane is entitled to affirm that *p*, I am initially excluding from my evaluation reasons that do not apply to Jane, in the sense that she does not see them as capable of motivating her conduct. But, if I applied this 'internalist' model rigorously,¹³ there would be no restriction at all, and Jane would be allowed to express any kind of complaint, using whatsoever kind of reasons, against the public justification of coercive norms. So, in order to even use the 'intelligibility' requirement meaningfully, those

¹³ (Williams, 1981) is the obvious example. If we had to use Bernard Williams's 'internalism' to public justification, we would arrive at the paradoxical conclusion that all 'internal' reasons, even those arrived at through mere delusion or deception, are usable in public justification.

that employ it must idealise their fellow citizens at least a little, so that they can adequately assess whether A can indeed affirm R_A . The problem is to see where this idealisation is supposed to stop.

An indication we receive from both Vallier and Gaus is that the idealisation, or, in other words, the move from real-life Jane to Member of the Public Jane*, must be minimal. I guess the final version of the idealisation, in their idea, must be something that the subject herself would be happy to accept, maybe adding the canonical phrase 'if only I had known better...'. This, at least, seems to me the most straightforward way to interpret Vallier's remark that '[somebody] is represented among members of the public as an idealized version of himself, one that only affirms reasons based on *sound information* and *valid inference*' (p. 604. My emphasis).

The problem is that the phrase 'if only I had known better..' can be uttered in all types of scenarios, including those in which the person has not merely tidied up her thoughts a bit, but has substantially changed her mind. Imagine our real-life friend Jane is a devoted Catholic who believes, among other things, that a) abortion is equivalent to murder, b) homosexuality is sinful, and, albeit tolerable, ought never to receive any form of public recognition and c) the theory of evolution ought to be taught at school only alongside intelligent design and divine creation. Now, of these three statements, maybe only the first one cannot directly be challenged by producing 'sound information' and 'valid inference'.¹⁴ Statement b) can be challenged through 'valid inference', by for example letting the person observe that the public recognition of homosexuality (say, through the legal establishment of same-sex mar-

¹⁴ One could advance the empirical evidence according to which fetuses begin feeling pain only at a very late stage of pregnancy and never really develop a conception of the self. But that would lead to debates about whether 'murdering' is equivalent to taking the life of someone who can at least feel pain, which is controversial and probably wrong.

riage) does not in any way threaten the sacredness she attributes to religious marriage; it is simply an alternative opportunity offered by the state. One could add that she herself recognises that state and religion have different functions and roles in public life, as she does not go to the town hall to pray or to church to renew her driving licence. Against c), one can claim that, once a theory has entered the scientific consensus, then it really makes no sense to teach it at school alongside previous ideas about the same matter, otherwise science teacher could not even start explaining classical dynamics without having first presented (as a plausible alternative) Aristotle's *Physics*.¹⁵

Now, imagine that, once confronted with this body of evidence, Jane would change her mind and utter the canonical phrase, 'if only I had known better...'. Shall we say that we are now dealing with a better version, epistemically speaking, of Jane, who has now reached the Gaussian status of Member of the Public (Jane*), or just that she is now a completely different type of reasoner (and is therefore \sim Jane)? I suspect both Gaus and Vallier would strongly suggest the second option, but I cannot see the reason why the first option is to be excluded. Jane, like it or not, was indeed failing to give due weight to relevant considerations when she retained her view that both a), b) and c) are valid statements.

Advocates of 'justificatory liberalism' could probably reply that I am now completely misrepresenting their project. I am trying to show that Jane has good 'external' reasons to reject her commitments to a), b) and c), whereas they want to consider only her 'internal' reasons, at least when they are not logically incoherent and obtain in situations where the agent is capable of reasoning (she is not drunk, drugged, hallucinating, etc....). But that opens up another epistemological problem,

¹⁵ I recognise this is of course a simplification. Intelligent design is, unlike Aristotle's physics, an idea that is supposed to be perfectly compatible with the rest of the contemporary scientific framework.

the difficulty of epistemic assessments across mutually incompatible viewpoints. Imagine Jane advances b) as a public view in order to show that the establishment of same-sex marriage is not justifiable to her. How can Mark, another member of the public who is an outspoken atheist, adjudicate whether Member of the Public Jane* is epistemically entitled to affirm b), according to real-life Jane's evaluative standards? Mark not only does not share Jane's evaluative standards but believes they are terribly wrong; whatever judgement he will produce about the way Jane is making use of her own standards is going to be grossly biased.

For Mark, judging whether Jane (or, for what matters, Jane*) is using her evaluative standards correctly is similar to judging whether someone who believed that triangles have four sides would then be epistemically entitled to affirm that squares have five. The 'assumption' that triangles have three sides is not one that we can either accept or reject so that, once it is clear which side of the debate you are on, I can still judge whether your reasoning, starting from that assumption, is valid. Once you start thinking that triangles have four sides, then my view of geometry and yours are simply incompatible, and, following from that, neither of us could act as a good judge in assessing the epistemic plausibility of each other's view.

Hence, at least on an epistemological viewpoint, the 'reasonable consensus' view seems much more attractive than 'convergence'. On the 'consensus' view, showing that a particular norm or institutional scheme is publicly justified is relatively easy, as it purely amounts to showing that it is indeed a norm or institutional scheme that reasonable people, suitably qualified, would agree on. As I know that reasonable people share some basic values and reason in a certain way publicly (i.e., they pay respect to the fact of reasonable pluralism), determining whether a consensus exists is not too hard.

On the ‘convergence’ view, by contrast, judging whether the reasons offered by citizens ‘converge’ on a commonly accepted scheme of public justification requires first assessing whether fellow citizens are entitled to affirm the views they are currently advancing, and secondly whether the latter either cohere or conflict with the justification offered by the state in defence of a certain norm or institution. Both passages suffer from the epistemic difficulty I illustrated before. My assessment of whether fellow members of the public are entitled to affirm the view they publicly advance, starting from their evaluative standards, is made virtually impossible by the fact that those evaluative standards might be stranger to me. Assessing whether my fellow member of the public is advancing a complaint against the enactment of a public norm that undermines its public justification is equally hard, because, as I do not share my fellow citizen’s evaluative standards, I can hardly judge whether their views constitute a decisive ‘defeater’ against an attempt at public justification offered by the state.

My argument is modest, as it points to a mere empirical difficulty, albeit of a particularly evident kind, and not to a logical impossibility, in assessing the epistemic plausibility of views based on assumptions one does not share. However, the empirical difficulty introduces the second problem I see with the ‘convergence’ model.

2.2.2 Convergence and the Fetishism of Public Justification

The ‘convergence’ view of political liberalism requires public officials to engage in significant investigations concerning the internal structure of conceptions of the good held by citizens, in order to see whether the citizens who publicly advance arguments based on them are epistemically entitled to do so. This requirement creates a burden for both the public authority and fellow citizens, who presumably will have

to wait till all complaints coming from civil society have been assessed (i.e., shown to be ‘intelligible’ or not) before they can see a law finally approved. Notice that those that advance a complaint, in the ‘convergence’ model, do so precisely because they prefer having no law on the matter than the law that is currently proposed.¹⁶ How can this burden be justified? Here we get to the core of the problem I see with the ‘convergence’ model of political liberalism, which is, to put it crudely, that it seems to cultivate a fetishistic interest in public justification, without disclosing the moral value underlying that interest.

In order to understand why public justification morally matters, we need a theory. My supposition, starting from the Rawlsian model of political liberalism, was that public justification mattered morally as a possible solution to the question of political legitimacy. This supposition seems to make less sense, however, once we move to the ‘convergence’ model. In case we used the ‘convergence’ model of public justification as a solution to the question of political legitimacy, the result we would achieve would be a slightly modified version of John Simmons’s neo-Lockean voluntarist model.¹⁷ Instead of letting the legitimacy of the state depend on the expressed consent of the citizens, we let it depend on the fact that the public justification offered by the state is not defeated by ‘intelligible’ complaints deriving from some citizens’ conceptions of the good. The result we are going to attain in both cases is presumably the same: a good argument in favour of anarchism. In terms of political ob-

¹⁶ More specifically, for Gaus there is a right against coercion that is violated every time a state enacts a new coercive law, and the law is not justified. Hence, ‘given our right not to be coerced, if the state wishes to enact law L, which would move us to [state] y, but it cannot justify φ -ing [coercing], it cannot justifiably move us to y. But if the state does not enact any law, we will stay at [state] x. So failure to justify can keep us at x’ (Gaus, 2011 b, pp. 487-488), which is the status quo. Or, in Andrew Lister’s (critical) words, ‘state action is justified only if not reasonably rejectable; if the cases for and against action are inconclusive, we default to inaction’ (Lister, 2010, p. 153).

¹⁷ (Simmons, 1999) (Simmons, 2015).

ligations, we would have to conclude that the state can only impose political obligations if the justifications for these obligations is not defeated by 'intelligible' complaints, which is an alternative way of stating that the state can impose political obligations only if citizens (reasonably idealised) agree with them.

As 'justificatory liberals' explicitly reject anarchism (Vallier, 2016 a, pp. 611-613), it follows that they do not conceive of public justification as a test for the legitimacy of the state. Their ultimate moral concern seems indeed different from that underlying the 'consensus' version of political liberalism. We can see this through Gaus's presentation of the 'problem of public reason'.

We thus arrive at the justificatory question: how can we identify social demands that all have sufficient reason to acknowledge as moral demands? This is the problem of public reason. (Gaus, 2011 b, p. 262)

Gaus is clearly addressing a vaster problem than the legitimacy of public institutions; he is considering, instead, the legitimacy of coercion in general, here understood as the legitimate imposition on others of 'social demands' that can be recognised as 'moral demands'.

I do not intend to discuss Gaus's interpretation of morality as a system of social demands that we need to justify each other. I want to notice, however, that it does presuppose two things. The first is a kind of overarching libertarian moral principle that, unlike the rest of morality, does not seem to have a 'social' nature, and as such is not itself subject to the 'public justification' test. The principle holds that

every person possesses an equal (natural?) right to reject any 'social demand' imposed on him or her.¹⁸ The second presupposition, usually much more explicit than the first, is the idea that the Members of the Public who are trying to justify to each other 'social norms' already possess pre-formed and somewhat complete 'evaluative standards' (Gaus, 2011 b, p. 276) and that they are going to use these in their acceptance or rejection of such norms.

The first presupposition is a controversial moral view, an aspect that Gaus seems to neglect and that leads to a contradiction in his theory. We cannot hold at the same time that all morality is social and deliberative, and hence subject to the possible rejection of people adhering to different 'evaluative standards', and that the libertarian presupposition in favour of liberty is undisputable. Some authors have tried to respond to this objection, by arguing that, insofar as the libertarian principle is not directly coercive, then it is not in itself in need of justification.¹⁹ But, if the problem for which we require public justification is that of direct coercion (the direct imposition on others of rules of conduct that they have to obey), then we would have to exclude from the test of public justifications all components of morality that do not lead to immediate coercive implications.

Furthermore, the principle is undeniably coercive in its implications, as it constitutes the main ground for accepting Gaus's version of public justification, which, in imposing on agents to justify to each other's evaluative system all moral rules, is one of the most burdensome systems of rules I can imagine.

In sum, the only distinctively moral principle that seems to underlie the model of public justification prevalent within 'justificatory liberalism' is one which,

¹⁸ In Gaus's own words, 'blameless liberty is the default: a person must have reason to endorse the authority of a moral rule' (Gaus, 2011 b, p. 319).

¹⁹ This is the canonical answer to the 'reflexivity' objection to public justification that we can find, e.g., in (Billingham, 2017).

when subject to the test of public justification, could be very plausibly rejected because in contrast with some member of the public's 'evaluative standards'. Hence, my allegation of fetishism towards public justification expressed before: 'justificatory liberalism' does not provide us with a moral reason for why we should adopt the particularly burdensome 'convergence' conception of public justification. In the absence of that, we are merely left with the impression that public justification is a kind of moral axiom on which the rest of the moral theory is construed.

In Chapter 4, I will offer 'justificatory liberalism' one last chance. I will try to show whether the principle of equal respect could represent an alternative basis for public justification *à la* Gaus/Vallier, but conclude that, if it were so, it would lead to even more implausible implications. Deprived of this last anchor, I will then argue that 'justificatory liberalism' ought to be rejected as an implausible theory.

2.3 Reasonable Consensus and the Problem of Interpretation

2.3.1 Preliminary Observations about Reasonableness

I will now head back to the Rawlsian version of public justification as 'reasonable consensus'. My aim is to prove that this model too fails to provide a solution to the question of political legitimacy. The discussion of 'consensus' will continue in the next chapter, where I am going to question the notion of reasonableness.

In the present section, I will focus critically on the other term composing the 'reasonable consensus' solution, that is, consensus itself. I will firstly give a reconstruction of what a reasonable consensus might plausibly be about, and then demonstrate that we cannot easily derive from it a solution to the question of political legitimacy. My objection, to anticipate briefly, is that a consensus about the reasonableness of certain abstract principles of justice does not entail a consensus about the reasonableness of all the infinite judgements that can be justified appealing to those

principles. It follows that citizens who agree about the reasonableness of certain principles justifying the authority of the state might still disagree about whether norms that are justified appealing to those general principles are themselves reasonable. And, as legitimacy, I will argue, can only depend on a combined and weighted evaluation of both the general principles justifying the authority of the state and at least some of the specific norms citizens are ordinarily subject to, a lack of consensus about the reasonableness of ordinary norms constitutes a decisive obstacle in the achievement of legitimacy.

Let's define a reasonable consensus as a consensus among reasonable agents about the acceptability of certain fundamental principles of justice and their use in the justification of political authority. The 'overlapping consensus' mentioned above is Rawls's own model of a reasonable consensus, but the argument I am going to offer does not necessarily depend on the consensus coming about through some overlap between the comprehensive conceptions of the good present in society.

What matters, on the other hand, is the reason why we want to let a solution to the question of political legitimacy depend on the existence and persistence of a consensus about general principles of justice. A Rawlsian response comes, I believe, from Rawls's definition of reasonableness.

The account of reasonableness that prevails within 'public justification' is one that bears a partial but important resemblance to ordinary language. In ordinary language, reasonableness denotes an attitude of modesty and considerateness for others' needs and interests both in the imposition of commands and the expression of judgements of a normative or evaluative kind. I am reasonable if I do not complain with my neighbours who were noisy until late yesterday night because I accept that they might want to celebrate the end of their exams. Similarly, I am reasonable when I do not mark students down when they overcome the word limit of their essay by a

couple of words, although, were I to decide differently, I would have the university rules on my side. In both cases, my attitude is reasonable, and not simply tolerant or, worse, submissive because my decision to abstain from a strict application of the rules was based on my understanding that particular features of the situations or the agents involved constituted valid excuses for their otherwise unacceptable behaviour.²⁰

Although the sense of reasonableness in ordinary language constitutes a preliminary model for Rawls, his use of reasonableness starting from *Political Liberalism* acquires a technical meaning that departs from ordinary usage. Specifically, reasonableness is associated with the possession of both strictly moral and moral/epistemic features.²¹ The strictly moral feature correspond, to reiterate, to an interpretation of society as a system of social cooperation for which one is only willing to propose 'fair terms' that he or she presumes others can accept (Rawls, 2005 [1993], pp. 49-50). Despite some similarities, the strictly moral features denoting the possession of reasonableness in Rawls bring to the fore a sense of reciprocity that is completely absent in ordinary language. In the Rawlsian sense, I behave reasonably in the public sphere when I do not propose principles that I believe others have good reasons to reject, provided they will behave likewise. In order to behave reasonably in the ordinary language sense, by contrast, I do not need to ask myself whether the rules I am imposing on others are acceptable to them. In fact, the pedantic professor

²⁰ Scanlon seems to stick more to the ordinary language of reasonableness when he explains the use of the adverb 'reasonably' in his contractual formula of morality in the following way: 'The intended force of the qualification "reasonably" [...] is to exclude rejections that would be unreasonable given the aim of finding principles which could be the basis of informed, unforced general agreement. Given this aim, it would be unreasonable, for example, to reject a principle because it imposed a burden on you when every alternative principle would impose much greater burdens on others (Scanlon, 2003, p. 132).

²¹ A similar distinction between moral and epistemological components of reasonableness is in (McMahon, 2014).

who marks down students who pass the word limit by a single word assumes that the students, as well as everyone else, either accept the norms as perfectly valid or, if they don't, are wrong. What distinguishes the reasonable from the unreasonable in ordinary usage is the ability to employ flexibility, grounded in consideration for others' interests and needs, in the application of rules. Notice that the flexibility that prompts me to tolerate the student who writes an overlong essay or my neighbours who loudly celebrated their last exam is a purely unilateral matter; which type of consideration I am going to accord to others' needs and interests is up to me.

The epistemic-moral features characterising reasonableness, by contrast, do not bear even a superficial resemblance to the ordinary language. Reasonable people, in the second sense, are defined by Rawls as those that acknowledge the role that the 'burdens of judgement' had in the genesis of reasonable pluralism and derive from that recognition a commitment to the 'duty of civility' and public reason (PL, p. 54). Rawls defines the 'burdens of judgement' as those 'hazards involved in the correct (and conscientious) exercise of our powers of reason' (PL; p. 56) that prevent a condition of reasonable agreement from becoming stable, at least on matters of value. Acknowledging the role of the burdens in the genesis of reasonable disagreement amounts to acknowledging that others disagreeing with us (again, on matters of value!) is not an immediate indication that they fundamentally misunderstood the matter (or, depending on our level of self-confidence, that we did). It might instead be an indication that disagreement on the matter will survive even the most scrupulous analysis and best efforts at reconciliation from both parts.

The moral-epistemic features characterising reasonableness create a substantive divergence from the ordinary usage. Indeed, in ordinary language we do not consider as a requirement of reasonableness the acknowledgement that one's own beliefs, despite maybe the most plausible, are just as reasonable as anybody else's.

We can observe that the ordinary-language sense of reasonableness possesses an uncontroversial moral character. Reasonable people are morally praiseworthy, in the ordinary-language sense, because they show to pay particular attention to others' circumstances when they behave morally, even when this leads to compromising principles of moral conduct that they generally commit to. On a deontological perspective, this attitude might be defended as one way (allegedly, among others) to discharge the fundamental moral duty of benevolence.²² By using flexibility in the rules we impose on others, a flexibility that is in turn grounded in an attentive consideration for the circumstances in which they behaved, we show ourselves to be benevolent.

I doubt that the moral value of the technical, Rawlsian sense of reasonableness is as straightforward. There are moral components that certainly characterise both the ordinary-language and the Rawlsian version of reasonableness, and these are more obviously recognisable as morally praiseworthy. But, even among the purely moral features characterising Rawlsian reasonableness, we find some concerns that, once properly analysed, appear morally controversial.

Take, for instance, the 'reasonable avoidance of the controversial' principle (RAC) demanding that citizens only advance in public terms of cooperation that they assume are acceptable for others. What is the moral value underlying this? The Rawlsian response is well-known: the 'political' interpretation of others in society as free and equal, which leads to respecting their 'political' autonomy (possession of

²² Beneficence (which, for the purposes of the present discussion, we take as equivalent to 'benevolence') is recognised by Kant among the 'duties of love to other human beings', defined as the 'maxim of making others' happiness one's end', whose duty 'consists in the subject's being constrained by his reason to adopt this maxim as a universal law' (Kant, 1996 [1797], p. 571). Sidgwick, unlike Kant, does not recognise benevolence as a duty, but as the 'supreme and architectonic virtue, comprehending and summing up all the others, and fitting to regulate them and determine their proper limits and mutual relations' (Sidgwick, 1884, p. 236). William David Ross, on the other hand, recognises 'beneficence' as a *prima facie* duty in (Ross, 2002 [1930], p. 21).

the capacity to form and revise a conception of the good and a sense of justice). But the jump from the understanding of others' freedom and equality to the requirement of 'avoidance of the controversial' is a significant one.

To understand why, compare RAC with another apparently similar moral requirement, which demands that we should not forcefully impose on others compliance with rules they believe they have no reason to respect, unless other conditions obtain. We define this as the principle of 'reasonable respect for autonomy' (RRA). Imagine I am a treasurer of an association and I have come to the conviction that a significant part of the social funds should be devoted to charity; alas, my partners disagree. Can I decide, unilaterally, to still devote the sum I had in mind to charity, despite my partners' contrariety? The moral answer is no, because, if I did that, I would be imposing on others the respect of a particular moral rule they believe they have no moral reason to respect. As I respect others' autonomy, I should refrain from exploiting my power (I would add regardless of whether, as in this case, my power derives from their consent) to impose on them the respect of those maxims of conduct that derive from my sectarian evaluative system. The principle is necessarily vague, because there seems to be lots of 'other conditions' that might obtain and that nullify the principle's validity, or that at least excuse its possible violation (for example when forcefully imposing a particular conduct on others is the only way to prevent a third party from being harmed), but it seems a robust moral principle nonetheless.

Notice, however, that the difference between RRA and the Rawlsian RAC lies in the verb 'impose', that comes to replace the Rawlsian 'propose'. The principle of reasonable respect for autonomy has a direct correlation with the exploitation of power. RAC *de facto* holds that moral agents who find themselves in a condition where they can impose rules of conduct on others should not exploit that power to

force others into adhering to their sectarian moral values. But, once we move from an asymmetrical to a symmetrical situation, as in the deliberative scenario Rawls is considering, the principle of reasonable respect for autonomy stops being applicable, for the very simple reason that no one can, in those conditions, *directly* impose the respect on others of sectarian moral rules, unless at least a majority of others consent.

I guess the intuition behind the Rawlsian RAC principle derives precisely from the case where a majority, following a deliberative process, imposes on a minority the respect of its sectarian moral values. That seems indeed unacceptable. We define this as the 'original intuition' of political liberalism. Now, what plausibly creates the unease in the 'original intuition' is the fact that the sectarian values were indeed *imposed*, so that the majority is now in the very same condition as the treasurer in the previous example, not the fact that they were previously proposed despite the fact that the minority could not accept them.

A Rawlsian could argue that one follows from the other, and that, if sectarian values cannot be imposed, purely because they are sectarian, then they should not be even suggested as terms of cooperation at the deliberative phase. But here we encounter a problem. 'Sectarian' is a purely descriptive label; it denotes those values or commitments that are shared in a particular community. Now, if we understand 'community' in a wide sense, meaning any group of people that shares some values or commitments, then it seems that every time we impose a rule in society, following a deliberative procedure of whatsoever kind, we are indeed imposing a sectarian view, namely the sectarian view shared by the party that prevailed in the process. Unless we consider anarchism a serious option, we have to readjust the 'original intuition' to take this into account.

Maybe, the moral problem that the 'original intuition' signals has nothing to do with sectarianism in general, but only with the imposition on others of values that have a particular historical pedigree and that are controversial within society. Following Rawls himself, we define a set of values and ideas that have a unitary structure and a recognisable history as a (comprehensive) doctrine. The moral problem in the 'original intuition' can then concern the imposition on others of norms of conduct that derive from specific doctrinal affiliations.

2.3.2 Reasonable Consensus and the Problem of Interpretive Disagreement

Can we make sure that a reasonable consensus avoids the problem of doctrinal impositions? Can we plausibly claim, in other words, that the moral values on which holders of reasonable conceptions of the good reach a consensus are not themselves part of a comprehensive doctrine? I fear many would say no, as respect for others' capacity to form and revise a conception of the good and a sense of justice might well amount to smuggling in some 'Kantian metaphysics' in a discourse which is supposed to be devoid of any controversial metaphysical commitment.²³ However, I do not intend to develop this critique.

My critique starts instead from the acknowledgement that the values and moral principles that can be part of a reasonable consensus compose a very limited set. In fact, they reduce to those moral and moral/epistemic considerations that are part of the definition of reasonableness itself. Any justification that relies on these values can seem acceptable to reasonable people; what else?

Assume there is indeed a consensus, within a single society, on a conception of social justice that the state appeals to in justifying its authority. For this to happen,

²³ For scepticism about the possibility of a value-neutral justification in political theory, see in particular (Hampton, 1989), (Bellamy & Hollis, 1995) or (Frazer & Lacey, 1995).

groups of citizens a, b, and c, all reasonable, must have reached an agreement on a specific conception of social justice, including 'constitutional essentials' P, Q, and R. Once they have agreed on theory of justice J {P; Q; R}, however, the normal working of politics has not even started. Indeed, everyday politics does not concern principles P, Q, or R, which we are assuming are part of the overlapping consensus, but specific norms that are justified appealing, at least in part, to their being consistent with P, Q, and R.

Hence, the complete scheme of public justification is the following. Reasonable citizens are each required to comply with norm n, which is justified at least partly appealing to principles P, Q, and R, together forming theory of justice J, which citizens have agreed on, in the Rawlsian model, through an overlapping consensus. The consensus was made possible by the fact that, *qua* reasonable people, a, b, and c already held a commitment to certain basic values, which are the moral and moral/epistemic components of reasonableness. As I wrote before that the moral and moral/epistemic components of reasonableness compose a basic conception of public morality, as opposed to a basic conception of justice, we call the set of reasonable requirements M {E; F} (with E and F the moral and moral/epistemic requirements respectively).

Now, let's assume that reasonable people do maintain their consensus on J stably enough not to compromise their allegiance to the state that is employing constitutional essentials P, Q, and R to justify its authority. This still does not settle the discussion on political legitimacy. P, Q, and R do not indeed constitute political obligations but are merely the principles used to legitimise *in abstracto* the existence of political obligations. In the quasi-Razian model of authority I am employing, the state has authority (*de facto*) when citizens take the directives it issues as capable of conditioning their conduct, and is legitimate when it has a right to demand that they

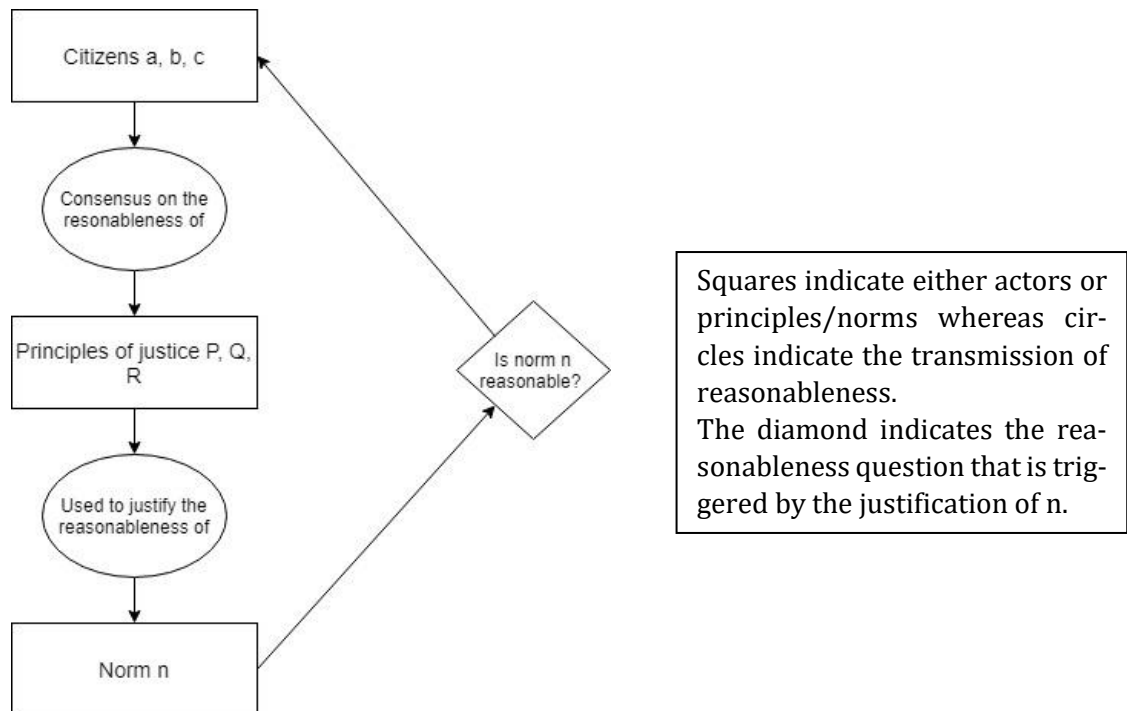
do so. When citizens consider whether the state is overall legitimate, therefore, they consider whether it is imposing on them a set of political obligations that is itself overall legitimate.

In order to evaluate whether the state is legitimate, therefore, citizens do not just consider the reasons P, Q, and R that the state uses to justify its authority in general, but combine that with an assessment of the entirety of norms included in the law, associated with the relative sanctions. They will need to consider, in other words, the rough sum of all the norms $\sum n$ ($n_1, n_2, \dots n_n$) that together form the overall pattern of compliance existing within a political society. I say 'rough sum' because I do not expect citizens to really evaluate (or, for what matters, even know) all the infinite norms, in all the various codes and statutes, that they are theoretically subject to.

These caveats notwithstanding, we can imagine that a positive evaluation of the 'rough sum' of the different normative requirements they are subjected to, combined with some equally rough acceptance of the general principles of justification P, Q, and R, is what moves citizens to consider the state ultimately legitimate. It is what moves them to adopt the characteristic conduct of pre-emptive adherence to norms we derived from the Razian model. Hence, the 'reasonable consensus' model would entail that reasonable citizens, in order to come to a consensus on the legitimacy of their state, ought to believe both that (1) the general terms of justification employed by the state to justify its authority are acceptable and (2) the pattern of compliance existing within that society (equivalent to the rough sum of all the significant norms) is also acceptable.

To form a judgement concerning the overall acceptability of the pattern, the reasonable agent will enquire on whether at least the most significant and burden-

some norms present in her society are broadly acceptable. Here is where the problems arise. For the reasonable agent, in assessing the acceptability of various norms, is supposed to engage in a further interpretation, regarding this time whether such norms are consistent with the reasonable theory of justice J that is presumed to justify the authority of the state. They are required to engage, therefore, in a double level of investigations regarding the reasonable acceptability of 1) the basic principles of justice that justify the authority of the state in the abstract and 2) the most significant and burdensome norms that directly demand their compliance. The analysis is summarised in the following scheme.



To see what the complexity of public justification implies in practice, let's assume that a majority of citizens has indeed reached a consensus on the acceptability of the constitutional principles of their society. Will that imply that they will also consider *every* norm that their political society decides to implement as reasonable? Of course not, as they might well believe that the drafters were mistaken when they believed that the norm in question was reasonable.

If possible, the original consensus on the acceptability of the constitution will give them even more reason to protest in case they believe a certain norm is unreasonable, as they can now claim that the norms is not only wrong, but betrays the spirit of reasonableness enshrined in the original constitution. This is an unsurprisingly common phenomenon in contemporary society; so common in fact that substantive debates are sometimes masqueraded as debates about the interpretations

of constitutional principles that every citizen is presumed, in view of the original consensus, to find reasonable. It is precisely through this logic that the NRA and advocates of the right to possess firearms in the US can portray themselves as defenders of the American Constitution, following a (legitimate, according to the Supreme Court²⁴) interpretation of the obscure Second Amendment. The reasoning behind is that, as long as there is a consensus on the reasonableness of the constitution, then any norm that is consistent with or even required by it (and in this case the Supreme Court itself claimed this was the case!) is also reasonably demanding citizens' compliance.

This reasoning is, however, both grossly unpersuasive and morally arbitrary. It is unpersuasive because no person who holds a firm conviction about the unreasonableness (in either the ordinary language or political liberal sense) of a particular norm could be possibly persuaded otherwise by the judgement that the norm is consistent with higher-order principles she herself generally accepts as reasonable. The person exposed to that judgement, even when it is coming from an authority she generally trusts, has at least two options that do not involve her changing her mind. She can firstly believe that the constitutional principle the norm refers to, although part of an overall reasonable system, is now so obsolete and archaic that can be easily dropped off from the system. Alternatively, and this is the most interesting case for the present discussion, she can believe that the court or whoever else defended the consistency between the norm and the constitution was mistaken and that the interpretation has turned a reasonable principle into an unreasonable request for compliance.

There seems to be nothing wrong in refusing to consider norm *n* reasonable, despite the fact that *n* has been justified by reference to either its consistency with

²⁴ *District of Columbia v. Heller*, 554 U.S. 570 (2008).

or derivation from higher-order principle P that one keeps considering reasonable. Reasonableness does not transfer from one higher-order normative principle to all the infinite lower-order norms that can be justified via reference to it, regardless of whether the justification itself is, according to necessarily controversial standards, correct.

It follows that the existence of a consensus on basic principles of justice not only is unable to alter the level of political disagreement that a given society contains, but creates its own form of *interpretive* disagreement. Advocates of reasonable consensus explicitly defend the view that these interpretive disagreements that take place within the reasonable consensus are different from ones that lie outside of it. My aim is now to show that this response is trivially correct but irrelevant once we take politics minimally seriously.

Jonathan Quong's distinction between 'justificatory' and 'foundational' disagreements is, in this sense, illuminating. In introducing the distinction, Quong presents to the reader two couples of friends (Quong, 2011, p. 207), who both debate what we may call substantive questions, namely the morality of using recreational drugs and the legitimacy for the Catholic Church of only hiring male priests. The second couple, unlike the first, shares a consensus about basic principles of social morality. When Tony and Sara, the protagonists of the second story, debate about the legitimacy of the Church only hiring male priests, their debate is constrained by 'public reasons' and they try to show that their view on the matter, rather than simply true or metaphysically correct, is the most 'reasonable', i.e. the one most plausibly derivable from the reasonable principles of public morality they both accept.

After having introduced us with the distinction, Quong draws some implications that seem far-fetched. He claims that ‘Sara can reasonably reject Tony’s position, and vice versa, but if the state were to act on the basis of Tony’s argument, it would be offering Sara an argument for the decision that she could reasonably be expected to endorse’. And that is because ‘Tony’s argument is one that Sara can understand and accept in her capacity as free and equal citizen, even if she does not believe it is the best argument, or even if she believes it to be incorrect’ (p. 209).

Once we relate this to the example before about gun control, we see that Quong’s conclusions do not stand. It is not the case that, because an argument is in principle accessible to me, in the sense that it is justified through values I consider reasonable, then I am bound to consider the outcome of that reasoning as maybe not the most reasonable, but reasonable nonetheless. I can still reject the reasoning itself as incorrect, as a mere attempt (malicious or benign) of including a reasonable value within a justification for a norm I consider deeply unjust. Hence, Sara can reject Tony’s reasoning in favour of granting autonomy in hiring to the Church by simply claiming that Tony is gravely misinterpreting freedom of religion, a ‘reasonable’ principle they are both committed to.

Is their ‘interpretive’ (or, in Quong’s words, ‘justificatory’) disagreement any different from the substantive debate of the first couple? If it is, it will be so in a way that is significantly devoid of practical implications. Indeed, assuming that in both cases one of the two parties is in the end able to implement the policy they advocated, the result will be identical, that is, the implementation of a norm that the other party has strong reasons to object. True, Tony in the second case provided arguments that tried to show Sara the reasonableness (as opposed to the necessarily ‘sectarian’ metaphysical truth) of the norm using principles that Sara herself is supposed to accept, but unfortunately for Tony, Sara believes that Tony failed in doing

that. Following from that conclusion, Sara now believes that allowing the Catholic Church to only hire male priests is *fully* unreasonable, despite the fact the state, through its attorney general Tony, has tried justifying that norm through a principle, freedom of religion, Sara finds reasonable and compelling. Sara might believe that Tony was sincere²⁵ in his attempt to use public reason to justify a norm she finds obnoxious, but she further believes that Tony is (inadvertently) subverting public reason.

Specifically, Sara disagrees on *interpreting* Tony's appeal to the reasonable principle of freedom of religion to justify the legitimacy of the Catholic Church only hiring male priests as a genuine instance of public reason. This disagreement entails that Sara is not persuaded the conclusions drawn by Tony are, even minimally, reasonable. In sum, an appeal to reasonable principles has generated a disagreement on whether the appeal in question was an interpretation of the principle that maintained, instead of corrupting, its reasonableness.

What are the implications of this problem for the legitimacy of the state? It proves that public justification based on reasonable consensus is incapable of reconciling political legitimacy and disagreement. We defined legitimacy as the right of an institution to demand that their citizens respect political obligations by adopting a certain attitude with regard to public norms, namely that of letting the institution pre-empt their conduct when norms are issued. Political liberals claim that the state acquires such right insofar as its authority is justified in terms that its subjects can reasonably accept. This way, according to their model, legitimacy is reconciled with the existence of disagreement.

²⁵ Quong takes sincerity of all the parties to be a condition of reasonable disagreement in Quong (2011), p. 207. Micah Schwartzman, on the other hand, takes sincerity to be 'part of an ideal of public reason' (Schwartzman, 2011, p. 377) and therefore, presumably, as demanded by the Rawlsian duty of civility.

But we also observed how, if we want to take disagreement seriously, we have to accept that citizens might have reason to contest both a *general* justification offered by the state to support its authority and a *particular* justification offered by the state to justify its requiring compliance for a single norm. And legitimacy cannot completely abstract from the particular norms that the state implements, because it makes perfect sense for citizens to retain their allegiance to the reasonable principles justifying the state's general authority and yet disagree about whether those same principles can be used to justify norms they strongly object. Legitimacy cannot completely abstract from the particular norms that the state implements, because the principles of justice object of the reasonable consensus (P, Q, and R in our scheme) could always be interpreted in the most disparate ways to justify the most disparate laws.

I tried to follow the suggestion, coming from political liberalism, that disagreement has to be taken seriously, in the sense that the state must justify its authority 'publicly' through reasons that citizens embracing different conceptions of the good can, at least if moderately idealised, accept. But the conclusions I reached are far less optimistic than the ones political liberals draw.

A possible attempt to rescue the political liberal approach to disagreement and legitimacy would consist in rejecting the model of legitimacy I exposed before, in which, in assessing the legitimacy of their institution, citizens consider not only the constitutional framework of their society, but also a 'rough sum' of the everyday norms to which they are subject. This is a theoretically sound move, but it would deprive the political liberal approach of some of its appeal.

Indeed, Rawls himself is not entirely consistent in his contention, inscribed in the liberal principle of legitimacy, that public justification only applies to 'constitutional essentials', or at the very least he does not seem to reduce 'constitutional

essentials' to constitutional principles. Instead, he mostly delimits the scope of public justification and public reason to the 'basic structure' of every political society, which 'is understood as the way in which the major social institutions fit together into one system, and how they assign fundamental rights and duties and shape the division of advantages that arises through social cooperation' (PL, p. 258). Significantly, Rawls lists among the elements of the basic structure, alongside the 'political constitution', the 'legally recognised forms of property, and the organisation of the economy, and the nature of the family' (ibid.).

I would say that my 'pattern of compliance' captures the spirit of the Rawlsian 'basic structure',²⁶ but, even if I were wrong, the contention that only constitutional principles matter for public justification is untenable. Due to the necessity of interpretation I mentioned above, two societies with identical constitutional principles can end up in opposite directions, in terms of everyday legislation. One interprets freedom of expression in such wide terms to forbid any limitations on the grounds of hate speech and slander, the other accepts that freedom of expression can be limited on the simple grounds of possibly causing distress to others. One interprets freedom of religion to be consistent with some forms of religious establishment (let's say the use of religious symbols in public spaces or official ceremonies), the other forbids that. One interprets the right to health to entail the creation of a freely accessible national health service, the other does not take a right to health to entail *free* access to health.

Shall we claim that they are both equally legitimate? What if a substantive minority in one of the two finds itself in Sara's position, in the sense that believes

²⁶ I tend to agree with those commentators who interpret the borders of the basic structure to be determined by the impact the basic structure has on citizens' welfare. See for example Tommie Shelby: '[I]t should be clear why Rawls chooses to focus on it: the effects of the basic structure on an individual's life prospects are immense and wide-ranging, and these effects have an impact on the quality of individuals' lives from the cradle to the grave' (Shelby, 2007, p. 129).

that some norms they are forced to comply with are not just wrong, but deeply unreasonable? I doubt that biting the bullet would be a compelling strategy for political liberals. My point is not challenging the distinction between justice and legitimacy. Instead, I am suggesting that a model of legitimacy that is sensitive to political disagreement must take into account the overall view citizens have of their state, and of the norms they are subject to, and not just their acceptance of some vague and variously interpretable constitutional principles.

To conclude; accepting, as in the political liberal paradigm, that societies acquire their legitimacy through some form of public justification accessible to mildly idealised people would lead to much more far-reaching consequences than those foreseen by political liberals themselves. It would entail that reasonable citizens can contest the legitimacy of the state not only when they believe that the general constitutional framework of their society is based on unreasonable principles, but also when they believe that, regardless of the constitutional order in play, some requirements of compliance that they are subjected to are unreasonable.

3. Reasonableness and the Problematic Epistemology of Political Liberalism

This chapter will end the discussion on legitimacy and public justification. It will focus on a single notion which plays, however, the most significant normative role in the literature on public justification, that of reasonableness.

In the previous chapter, I have concluded sceptically about whether we can let a solution to the problem of political legitimacy depend on a reasonable consensus about basic principles of justice. In the present chapter, I will move to the other term of the phrase 'reasonable consensus', that is, reasonableness.

In the previous chapter, I noticed that the moral value of the technical sense of reasonableness that Rawls introduces in *Political Liberalism* is far from obvious. When we move from ordinary language to the political liberal jargon, that is, a part of the intuitive moral appeal of reasonableness disappears.

This is not a decisive reason to abandon the Rawlsian notion of reasonableness. Rawls does not take reasonableness as a moral virtue, whose value must be transparent to our moral intuitions, but as a political attitude that we adopt with a specific purpose in mind, which is that of reaching a consensus on fundamental matters of justice with our co-citizens.

In the present chapter, however, I want to investigate further what adopting a Rawlsian attitude of reasonableness amounts to, especially on an epistemological perspective. When introducing the notion of reasonableness, Rawls seems happy to accept (or, better, to concede *in passim*, considering that the expression appears between brackets) that reasonableness 'has epistemological elements' (PL, p. 62), but he never elaborates on such elements, beyond the requirement that the reasonable

acknowledge the 'burdens of judgement' and their role in the emergence of reasonable disagreement.

But, if we want the notion of reasonableness to play a significant normative role in a theory of justice, both its implications and its underlying assumptions must be disclosed. My final reflections on this matter will be that reasonableness unfortunately does lead to implausible implications as it does commit either to a form of scepticism or to a very bizarre and possibly irrational epistemic attitude.

As I will specify later, the scepticism I am considering concerns the credibility of one's conception of the good, or, more precisely, of beliefs derived from one's conception of the good. My conclusions are going to be that committing to reasonableness demands from agents that they attribute a lower level of credibility to the beliefs of theirs that are the object of reasonable disagreement than to beliefs that are not affected by the same phenomenon. This conclusion does not amount to a full-blown scepticism regarding one's conception of the good, but it does rule out the possibility that, to exemplify, a reasonable religious believer might be as certain of some religious precepts as she is of scientific results that are not similarly object of reasonable disagreement.

This is a paradoxical result for political liberalism. Political liberalism's *raison d'être* was the elaboration of a theory of liberal legitimacy that paid respect to the phenomenon of value-disagreement. But political liberalism in fact demands that reasonable citizens subject their values and commitments to significant doubts. Once this has occurred, however, the disagreement persisting among reasonable citizens is much less of a threat to the liberal order.

As this is the last chapter on legitimacy and justification, I will conclude by drawing some implications that derive from my analysis of legitimacy and disagreement. I will argue that the question of legitimacy can never be reduced to the simple

public justification of the authority of the state as an account of legitimacy can never completely abstract from the subjective expression of approval or discontent that citizens offer regarding policies that particularly affect them.

However, these subjective expressions constitute raw data that must be both adequately weighted, as they do not compose a perfectly coherent pattern, and possibly filtered, as they might include all sorts of callous preferences that we might want to discount. I will conclude that the best solution to the problem of reconciling political legitimacy with value-disagreement is one that is capable of combining objective and subjective elements. The subjective elements consist of the various perceptions citizens hold regarding both the general justification of the authority of the state and the reasonableness of the main norms to which they are subject. The objective elements are certainly more problematic as they presuppose a comprehensive theory of morality that complements an account of legitimacy. Almost pointless to say, the present work does not constitute even an attempt at building that 'comprehensive theory', although the next two chapters might well be taken as components of a vaster project in that regard. Despite this difficulty, the inclusion of objective criteria within a theory of legitimacy is made necessary to prevent the conclusion that, because of the lack of directly expressed, unanimous, consent, most existing states are illegitimate.

This chapter will conclude, therefore, in slightly pessimistic tones. Not only reconciling legitimacy and disagreement is a project that still requires a significant effort to be completed, but the most prominent attempt at arriving at a solution ultimately fails. The next chapters will bear the burden of furnishing the constructive component of the present research. They will show how, leaving aside legitimacy, we can organise a good liberal response to the question of disagreement following instead the two normative principles of equal respect and toleration.

3.1 A Critique of the Appeal to Reasonableness in Public Justification

3.1.1 *Introducing the Problem*

I will begin with a personal anecdote. During a Christmas Eve dinner a few years ago, I had an interesting conversation with my ninety-year-old auntie regarding the controversial (for her) topic of same-sex marriage. I was pleased to hear she had no objection against the state instituting a scheme of civil partnerships for same-sex couples (as in the current Italian law), but, as a Catholic, she would never accept the state recognition of same-sex marriage. The motivations she advanced tell a lot about the idea of the relationship that, according to many, ought to exist between state and cultural conventions. She claimed that of course she recognised the distinction between church and state, but also pointed out that the state cannot completely subvert a noble tradition by suddenly deciding to call '*matrimonio*' ('marriage') something that could possibly not even include a *mater*.¹

Many would be inclined, I believe, to call out her behaviour as unreasonable. This allegation would probably cohere with ordinary language, at least among people who believe same-sex marriage is morally unproblematic. However, once we analyse the matter without prejudice, we see that the exact reason why my auntie's attitude qualifies as unreasonable, in the Rawlsian sense, is far from obvious.

My auntie was advancing considerations that were not purely based on her religious affiliation; her point about the state subverting a long tradition by changing one fundamental feature of marriage has nothing to do, at least in the way she presented it, with the condemnation of homosexuality as a sinful deed. It has more to

¹The use of Latin is still a kind of trump card in discussions in Italy, especially among people born in the first half of the last century. She seemed unimpressed by my sarcastic remark that we don't need a *pater* to have a '*patrimonio*' (Italian for 'capital' or 'endowment'). I hope the reader will concur that raising the more serious issue of lesbian *matres* would not have helped me win the argument.

do, I suspect, with her expression of a fundamental interest she had in the preservation of those features of the surrounding world she had witnessed since childhood and gradually learned to value. I do not intend to argue that this ‘conservative interest’² is part of everyone’s conception of the good, but only that it is an interest sufficiently accessible, intelligible, and independent from sectarian commitments to possibly figure in ‘public reason’.

This being the case, my auntie’s view about same-sex marriage cannot be merely dismissed as the expression of a sectarian preference whose inclusion in the public debate would display a scarce engagement with the ‘burdens of judgement’ and the constraints they originate. Moreover, in her accepting that the state might recognise civil partnership even for homosexual couples, she was showing at the same time respect for others’ freedom and equality and that she was willing to partially compromise her ‘conservative interest’. Following from this, a Rawlsian liberal would be probably more inclined to include a conservative view expressed in those terms within reasonable pluralism.

Here again we see, however, an important contrast with the ordinary sense of reasonableness. In the ordinary sense, I would not call an opinion reasonable, regardless of its content, just because it expresses an interest that is sufficiently accessible to the general public (in whatever way I decide to flesh out accessibility – whether as independence from sectarianism or ‘intelligibility’ or whatever else) to qualify as ‘public’. A further, and more serious, problem is that, if we follow the political liberal model, a recognition of another’s opinion as reasonable is a normatively salient occurrence. It entails a particular kind of public treatment that is owed

² Among the liberal thinkers who have recognised not merely the intelligibility, but the reasonableness, of that interest, are Joseph Raz in (Raz, 2001), Samuel Scheffler, with his fantastic remark that ‘if there is a conceptual gap between valuing and the impulse to conserve, it is not a very large one’ (Scheffler, 2007, p. 107), and, more than anyone, G.A. Cohen (Cohen, 2011).

to the person whose view has been recognised as reasonable, in the sense that her views count now among the ones that need to be addressed by the state in justifying its authority. Which in turn means, connecting this to the example, that the state, in justifying its official recognition of same-sex marriage,³ needs to address the conservative objection coming from citizens whose main point is, to put it crudely, that they do not want to live in a social environment that is too distant from the one they were used to.

Now we come to the central question, which is whether this use of reasonableness is, in the end, plausible. In particular, I intend to question whether reasonableness, in the Rawlsian sense, would entail some unattractive (for Rawlsians) form of scepticism or relativism. The question is made relevant by the fact that, were reasonableness to entail some form of moral scepticism or relativism, that would immediately destroy the appeal of the entire notion, as embracing reasonableness would require rejecting one's belief in the truth of one's conception of the good. Let alone that this is explicitly denied by Rawls,⁴ interpreting the overlapping consensus as a consensus of sceptical or relativistic doctrines would push the consensus even further away from the reality of political disagreement.

Unfortunately, despite Rawls's remarks on the matter, taking the 'burdens of judgement' seriously on an epistemological perspective does lead to either some form of suspension of judgement about the beliefs that are subject to reasonable

³ Following the model of *de facto* legitimacy I introduced at the end of the previous chapter, which encompasses both judgements regarding the general constitutional principles and the 'rough sum' of everyday legislation, I am here accepting that the domain of application of public justification can be expanded, *contra* Rawls, beyond constitutional essentials and matters of basic justice (see the limits of the application of public reason in PL, pp. 228-30). That the Rawlsian restriction makes sense is denied by both Gaus (see his discussion of 'what is to be justified' in (Gaus, 2011, pp. 490-497) and by the usually much more sympathetic Quong (Quong, 2011, pp. 273-287).

⁴ Rawls explicitly denies that either the notion of the burdens of judgement (PL, pp. 62-63) or the overlapping consensus (pp. 150-154) can lead to sceptical conclusions.

disagreement or to an implausible epistemological attitude. To recapitulate, Rawls argues in *Political Liberalism* that, when reasoning in certain specific domains (which Rawls does not exactly specify, but which we can take as involving value in general – whether moral or aesthetic), the agents involved are subject to some ‘burdens’ that prevent them from (indefinitely?) reaching an agreement. Rawls lists some of these burdens,⁵ but specifies the list is not meant to be ‘complete’ (PL, p. 56) – the appeal of the reasoning, therefore, is mainly intuitive.

The intuitive appeal of the ‘burdens’, however, cannot justify the particular normative requirement that Rawls derives from them, namely that we treat even views that we strongly disagree with, including my auntie’s, as constituting a possible constraint on the public justification of compliance-demanding norms. Why should I accept that views I find morally problematic, in virtue of their being ‘public’, may constitute a constraint on the public implementation of compliance-demanding norms? Merely pointing to the inevitability of disagreement is not enough. If I want to take the burdens of judgement seriously, however, or precisely seriously enough to lend support to demanding normative implications, I need to draw a different conclusion, namely that those I disagree with are as justified as I am in holding their views on the matter. If this is the case, however, shouldn’t I lower my confidence in the plausibility of my own view? This is the unwelcome conclusion that I believe political liberalism is inevitably committed to.

⁵ ‘a. The evidence – empirical and scientific – bearing on the case is conflicting and complex, and thus hard to assess and evaluate. b. Even when we agree fully about the kinds of considerations that are relevant, we might disagree about their weight [...]. c. [A]ll our concepts, and not only moral and political concepts are vague and subject to hard cases; and this indeterminacy means that we must rely on judgement and interpretation. [...] d. To some extent [...] the way we assess evidence and weigh moral and political values is shaped by our total experience [...]; and our total experiences must always differ. [...] e. Often there are different kinds of normative considerations of different sides on both sides of an issue and it is difficult to make an overall statement. f. Finally, [...] any system of social institutions is limited in the values it can admit so that some selection must be made from the full range of moral and political values that must be realised’ (PL, pp. 56-57).

3.1.2. Reasonableness and the Problem of Scepticism – Rejection of the Epistemic Systems Paradigm

The reasonable disagreement of political liberalism cannot be confused with the disagreement among epistemic peers that is discussed in social epistemology; the two are distinct because in reasonable disagreement the parties cannot recognise each other as having equal (or even roughly equal) access to the appropriate evidence. They cannot, because what counts as the ‘appropriate evidence’ that might justify their particular doxastic attitude in relation to a certain matter (whether belief, belief-that-not, suspension of judgement, or a graded form of any of these) is precisely a matter of dispute. Still, recognising somebody as a partner in a reasonable form of disagreement, which originated due to the effects of the ‘burdens of judgement’, must lead to some consequences regarding one’s approach to the belief around which the disagreement revolves.

The least troublesome (on an epistemological perspective) approach to reasonable disagreement is suggested by Alvin Goldman. It consists in allowing for a form of ‘objectivity-based relativism’ (Goldman, 2010, p. 196 ff.) according to which agents disagreeing on a certain matter are all justified in retaining their doxastic attitude, despite these pointing to opposite directions, because they respectively adhere to different epistemic systems, defined as ‘sets of norms, standards, or principles for forming beliefs and other doxastic states’ (p. 187). As long as they are all respecting the ‘norms, standards, or principles’ of their respective epistemic systems, they are, at least in a sense, objectively justified in holding their belief despite the disagreement. Goldman defines that particular sense of justifiedness ‘second-order justifiedness’, defined as ‘the status of being [objectively] justified in believing that they are [objectively] justified in believing P’ (p. 202). This form of relativism

does not contrast with epistemic objectivism because, even though at most only one of the epistemic systems is itself epistemically correct (in whatever sense we understand epistemic correctness), different agents might ‘occupy different evidential positions *vis-à-vis* [their] system and other candidate E-system. Hence, the objective justificational status of different people *vis-à-vis* different E-systems is varied rather than uniform’ (p. 201).

Goldman’s conclusions are relevant for the debate, but I would be more cautious than Fabienne Peter is in drawing implications from them that are broadly favourable to political liberalism,⁶ for at least two reasons. The first is that, in order for the parties to be all justified in holding their mutually contradictory beliefs, we need to keep assuming in reality what Goldman only assumes in theory,⁷ namely that in reasonable disagreement people all occupy different ‘evidential positions’ that make them incapable of judging the overall plausibility of the epistemic system they embrace. The second is that, if we really wanted to hold that assumption and conceive reasonable disagreement as an ensemble of agents each equally justified to hold their beliefs, then the distinction between reasonable disagreement and disagreement among epistemic peers would indeed falter, which will lead, however, to the sceptical implications.

⁶ See (Peter, 2013, p. 608 ff.). Similarly, Laura Valentini tries to derive a defence of liberal democracy from the recognition of reasonable pluralism, but the argument is, to me, similarly problematic. Valentini argues that ‘[d]emocracy is what equal respect (procedurally) requires when there is thick reasonable disagreement about what equal respect (substantively) requires’ (Valentini, 2012, p. 193). Applying the argument I am going to lay out to Valentini’s main point, if democracy is indeed the best solution to the existence of reasonable pluralism, then parties that accept democratic rules ought to accept that their own convictions that were subject to reasonable disagreement, and that generated the need for a democratic solution, ought to be questioned.

⁷ Goldman repeatedly claims that his presentation of the different epistemic systems is purely speculative.

Starting with the first point, the circumstances that would make the assumption of agents in ‘different evidential positions’ plausible are hard to obtain. An example provided by Goldman (pp. 198-199) can help explain why (*contra* him). In relation to the debate in the US regarding the teaching of evolutionary theory at school, Goldman writes, ‘Those instructed in a modern biology curriculum are taught to assign high credence to whatever evolutionary science says about the world. [...] By contrast, students taught in fundamentalist schools [...] are taught to be sceptical about whatever evolutionary science says’ (pp. 198-199). Now, a scenario of this kind does indeed produce two apparently different epistemic systems, namely E (‘whatever evolutionary science says is reliable’) and \sim E (‘whatever evolutionary science says is unreliable’).

But this representation is of course inaccurate because E and \sim E are not interpretable as epistemic systems, but, at most, as single epistemic norms within a more comprehensive epistemic system. In a more realistic picture, we would have two epistemic systems having the form E* (‘trust in general the scientific consensus + a number of other epistemic norms’) and \sim E* (‘trust in general the scientific consensus *except in the case of evolution* + a number of other epistemic norms’).⁸ If we take this picture as a fair representation of the debate, then individuals who respectively endorse E* or \sim E* are not locked within mutually incomparable epistemic systems; on the contrary, they share a good deal of epistemic norms with the other parties in the debate, a circumstance that is made virtually unavoidable by the co-existence in the same society. As they share some epistemic norms, they are in a good position (in fact, I would say, in the perfect position) to judge whether their

⁸An even more precise representation would probably assign some numerical value to the degree of credence that each party attributes to the scientific consensus in general. Thus we would have E*{75% credence on scientific consensus + other epistemic norms} and \sim E*{55% credence on scientific consensus + 0% credence on evolutionary science + other epistemic norms}.

own epistemic system, or somebody else's, is the most correct, according to standards that are sufficiently common (say, internal coherence, comprehensiveness, capacity to account effectively for phenomena in the real world etc.). Hence, it is untrue that disagreeing agents within the same society can occupy such different 'evidential positions' that they are epistemically justified to follow uncritically what their epistemic system demands them to do.

This argument should have proved the unlikelihood, in the context of contemporary societies, of scenarios where agents can be said to disagree reasonably because they are each entitled to both a) follow the epistemic norms of their epistemic systems and b) adopt a kind of uncritical approach to that epistemic system. Alas, somebody like Peter could object that, although this is true in general, reasonable disagreement (of the Goldmanian type) occurs specifically in the area of value-disagreement. The argument behind it would be that the various conceptions of the good are interpretable as epistemic systems that people follow when they need to achieve something-like-truth (epistemic correctness) in the realm of values. Some conceptions of the good, mainly established religions, work indeed like epistemic systems, insofar as they communicate to their followers that a specific source (a sacred book, the words of a priest, etc.) constitutes reliable evidence.

I am not sure that would equally apply to all conceptions of the good, as some conceptions of the good seem to communicate to their followers exactly *what* is good or valuable, and not what is the evidence to be considered in determining what is good or valuable.⁹ Still, there is a crucial distinction between value-disagreement

⁹ I guess someone could say that utilitarianism acts as a kind of epistemic system insofar as it does provide its followers with a heuristic strategy to determine what is right. I still believe that there is a difference between a mere heuristic device to derive a conception of the right (such as the utilitarian maximisation of aggregate utility or the Rawlsian original position) and a comprehensive epistemic system, but I will not press that particular point.

and factual disagreement in the sense that not all the parties involved in value-disagreement intend to question the plausibility of their own or somebody else's conception of the good.

To refer back to the original example in this section, my Christmas Eve dinner discussion did not revolve centrally around the question of which kind of evidence should people rely on in formulating a correct judgement about same-sex marriage. A more accurate characterisation would be as a discussion about what kinds of *things*, which presumably we all agree are valuable (at least for people in general, if not for ourselves) ought to be taken into account in formulating that judgement. As I argued before, the 'conservative interest' my auntie was advancing is not entirely stranger to me; certainly not so stranger that, in order to appreciate it, I would have to convert to a different epistemic system. On the contrary, it is an interest whose relevance I can perfectly understand, and that I simply choose not to prioritise in the particular debate I am conducting. I doubt, on the other hand, that we could have a debate at all, if we had both referred to evidence that was not, at least in part, accessible to each other, in the sense that we both agree it might have some relevance for the matter.¹⁰

3.1.3 Reasonableness and the Problem of Scepticism – Rejection of Extreme Permissivism in the Face of Reasonable Disagreement

I am sceptical, therefore, that value-disagreement can ever be characterised as disagreement based on the participants' reliance on different epistemic systems. However, were it indeed the case, what would be the epistemically correct attitude for

¹⁰ During the debate on civil partnership in the Italian Parliament, a conservative parliamentarian by the name of Domenico Scilipoti tried approaching the debate from a purely sectarian, Catholic perspective. He mentioned in his address to the Senate all the biblical verses that condemn homosexuality. Even the Italian Senate, not exactly a liberal temple, laughed at the use of that particular evidence in the debate.

people to adopt? The political liberal suggestion is that, following the acknowledgement that the disagreement is caused by the presence of different epistemic systems, none of which is in the current circumstances demonstrable to be superior, the parties can both:

- retain their belief that p (commitment to their conception of the good) and
- a) refrain from using p publicly, as p is part of a reasonable disagreement ('doctrinal avoidance') and
b) demand that political legitimacy only takes place through a pattern of justification that, because it does not contain p-like principles, is accessible to all epistemic systems ('doctrinal compatibility').

I will now argue that this suggestion leads either to a form of scepticism or to an irrational epistemic attitude. Parties who have reached the conclusion that a reasonable disagreement is in place believe that the truth of p, the belief they entertain, is evident to them but inaccessible to others who commit no fault by rejecting p. In one of the most sophisticated defences of reasonable pluralism,¹¹ this idea was defended by appealing to the assumption that disagreement on certain matters 'persist at all levels of competence' (Van Wietmarschen, forthcoming); no matter how idealised two epistemic actors are, they will never agree that p or non-p.

¹¹ (Leland & Van Wietmarschen, 2012). For an equally sophisticated critique, see (Enoch, 2017). The Rawlsian inspiration for Leland and Van Wietmarschen is from PL, p. 58: 'many of our most important judgements are made under conditions where it is not to be expected that conscientious persons with full power of reason, even after free discussion, will all arrive at the same conclusion'.

But, being this the case, the reasonable participants to the reasonable disagreement must necessarily lower their confidence in the belief that is disputed.¹² In order to defend this point, I intend to argue for a minimalist version of what epistemologists call a Conciliatory View in the face of peer disagreement.¹³ The view I now come to defend holds that people should accord a lower degree of confidence to a belief that is the object of peer or reasonable disagreement than to one that it is not.

Imagine I am with a friend and we are trying to recollect some memories from a journey to Tuscany some years ago. We both agree that we visited Florence, but only I remember visiting Siena. Now, unless I have reason to suspect my friend's memory is generally less reliable than mine is, I cannot hold with the same level of confidence the belief that we went to Florence and the belief that we went to Siena. I might still retain part of my confidence that we did, in the end, end up in Siena; after all, my memory is usually reliable evidence.¹⁴ However, the disagreement

¹² As Han Van Wietmarschen has recently admitted, in a piece that was a significant source of inspiration for these reflections (Van Wietmarschen, forthcoming).

¹³ The Conciliatory View is defended in (Christensen, 2007) (Elga, 2007) (Van Wietmarschen, 2013). David Christensen defines conciliationism as the view 'on which the disagreement of others should typically cause one to be much less confident in one's belief than one would be otherwise – at least when those others seem just as intelligent, well-informed, honest, free from bias, etc. as oneself' (Christensen, 2009, p. 756). Christensen further defends, however, a principle that he calls Independence and that he believes all holders of conciliatory views must accept: 'In evaluating the epistemic credentials of another person's belief about P, to determine how (if at all) to modify one's own belief about P, one should do so in a way that is independent of the reasoning behind one's own initial belief about P' (p. 758). As noticed by many critics (see especially Kelly 2013, pp. 37-43), there seem to be too many counter-examples to the idea that, in order to have a rational response to peer disagreement, you should discount all the evidence that brought you to believe that p when you try to determine the reliability of your opponents in the disagreement about p. Hence, the minimal conciliatory view I defend here does not commit to Christensen's Independence.

¹⁴ The view according to which one is entitled to attribute an 'extra weight' to his view is defended in (Enoch, 2010) and, to a certain extent, (Feldman, 2009). Even though advocates of Extra Weight usually emphasise its distinction from the conciliatory view (see footnote above), the two are, for the purpose of the present discussion, just two variants of a wider view, according to which disagreement with epistemic peers requires you to revise, in some sense, the reliance you place on your belief that is object of disagreement. Richard Feldman, for example, argues, against strict conciliationism, that there are no specific epistemic principles that apply to the case of disagreement but that, in general, '[t]he proposition that S's peer who shares S's evidence concerning P disbelieves P is evidence against

forces me to abandon some of my confidence in the belief that the visit to Siena did occur. If I refuse to accord a lesser degree of confidence to my belief that I went to Siena compared with that of my belief that I went to Florence, then I am indeed being irrational.

The previous example is supposed to support a very minimal conciliatory view, which in particular is silent on whether the agents have to substantially change their belief in p (for example moving from p to p^* , which is closer to the $\sim p$ expressed by the disagreeing agent) when p is the object of a peer disagreement. The agents are just required to lower their confidence in p , which, to put things a bit less cryptically, might correspond to an epistemic attitude such as ‘starting to cultivate some doubts regarding p ’.

The scenario described above about the disagreement in recollecting the various stops in a journey does not share all features in common with a case of disagreement (whether reasonable or not) about value, because, as I argued before, the very definition of competence and expertise is, in value-disagreement, either contested among the disagreeing agents or generally unclear. However, in order to avoid scepticism, it is not enough to point out that reasonable value-disagreement is hardly equivalent to peer disagreement because the definition of competence is contested and is therefore hard to say whether the participants to the debate are effectively epistemic peers. To retain full confidence in one’s belief, one would also need to possess some scepticism-preventing epistemic warranty that one’s convictions on the matter are superior to anyone else’s.¹⁵ The presence of such warranty,

P’ (p. 298). Similarly, according to Thomas Kelly’s Total Weight View (Kelly, 2010; Kelly, 2013), disagreement is just one piece of evidence to be considered, among the others the subject has at its disposal, in forming a final decision regarding the credibility of P.

¹⁵ Similar reflection in (Simpson, 2013, p. 572): ‘Parties to a dispute about any complex issue are unlikely to be strict evidential and cognitive equals in relation to the things that they disagree about,

however, would not merely relieve some of the agents involved (those that are protected by the warranty) of the epistemic duty to ‘start to cultivate some doubts’ regarding their disagreed upon value-convictions. It would also eliminate the epistemically reasonable character of the disagreement, because the disagreement could not be characterised any longer, following Leland and Van Wietmarschen again, as one in which disagreement can persist at all levels of idealised competence; if there are scepticism-preventing warranties, then competent agents will be able to identify them sooner or later.¹⁶

Some epistemologists contest the conciliatory view about peer disagreement; such supporters of Epistemic Permissiveness usually reject the Uniqueness View, according to which, ‘for any body of evidence E, and proposition P, there is only one doxastic attitude to take towards P that is consistent with being rational and having evidence E’ (Schoenfield, 2014, p. 195). The conciliatory view I have defended so far is, however, very modest and purely applies in comparative cases, where one agent is required to assign a different level of credibility to two or more of the beliefs he entertains. In case one such belief is subject to reasonable disagreement and the other is not, I argue that there is indeed only one doxastic attitude for the agent that is rational to take:

but it does not follow that either one of them can justifiably downgrade the other from the status of epistemic peer to epistemic inferior.’

¹⁶ Leland and Van Wietmarschen argue that reasonableness includes a demand of modesty that prevents reasonable people from idealising too much when they need to determine whether a certain position is still disputed at high levels of competence. Their solution is that the reasonable ‘is permitted to idealise toward any level of competence *present among his reasonable fellows*’ (Leland & Van Wietmarschen, 2012, p. 731 – my emphasis). But this ‘modesty in expectations’ is accompanied by ‘intellectual modesty’ that demands that ‘[f]or each of his nonpublic convictions and political conclusions, [the reasonable] believes that reasonable people at all levels of competence endorse views that conflict with his own’ (ibid.). Hence, even though reasonable disagreement in their view need not always involve matters that are disputed at the highest possible levels of competence, disagreement about matters of value need to be characterised (because of ‘intellectual modesty’) in exactly those terms.

For any two beliefs P and Q, if P is subject to reasonable disagreement and Q is not, then, *all other things being equal*, the rational doxastic response for the agent is to assign P lower credibility than Q. (Modest Uniqueness)¹⁷

The question of import here is not whether Modest Uniqueness is all-things-considered correct (although I do believe it is very highly plausible), but whether it is convenient for political liberals to reject it. I think it is not, for a double set of reasons.

Firstly, the conciliatory epistemic attitude in the face of (peer or reasonable) disagreement has a significant role to play in the evolution of my knowledge. Lowering my confidence on p, when p is reasonably disagreed, will prompt me to look for new evidence about p, until the point I am no longer an epistemic peer of my friend, but an epistemic superior (which will happen if, following the previous example, I finally manage to find a photo of our visit to Siena). Defenders of Extreme Permissiveness, who deny Modest Uniqueness, would have to accept that their view deprives agents of this important epistemic advantage and of a pragmatic use, in everyone's favour, of the existence of disagreement.

¹⁷ Some important clarifications. In Modest Uniqueness (MU), I am considering beliefs that the subject already holds, rather than proposition he has to decide whether to believe or not. This is an important distinction between MU and Uniqueness because MU is silent on whether subjects can still be justified in believing (though at a lower level of confidence) a proposition that is subject to reasonable disagreement. This shields MU from some important objections, for example Miriam Schoenfield's point that 'certain plausible theories of justification require the truth of permissivism' (Schoenfield, 2014, p. 197). As I wrote, MU differs from Schoenfield's version of Uniqueness insofar as it does not exclude that different agents might be justified in believing opposite propositions to be true, but only that, in case that happens, and they recognise that the disagreement is reasonable, they must lower their confidence.

Secondly, the specific version of Uniqueness I have so far suggested seems to be entailed by the recognition itself of reasonable disagreement. The very talk of a reasonable disagreement implies that there are (at least) two forms of disagreement, or, to follow the Rawlsian jargon, two forms of ‘pluralism’¹⁸ (one reasonable and the other non-reasonable) and consequently two types of challenges that disagreement can pose towards an agent’s belief in the truth of a certain proposition. It seems indeed perfectly consistent with, if not implied by, the recognition of the role of the burdens of judgement (one of Rawls’s conditions for being reasonable!) to assign lower credibility to the beliefs that are affected by the burdens *vis-à-vis* those that are not.

I do not intend to add anything to this minimalist defence of a conciliatory view in the face of disagreement; defenders of Extreme Epistemic Permissiveness might be unpersuaded, but I doubt political liberals would be happy to bite the bullet here and align themselves with such a controversial epistemological view.¹⁹

If, however, Extreme Epistemic Permissiveness is ruled out, agents recognising that certain beliefs of theirs are part of a reasonable disagreement (and who want to maintain rationality) must lower their confidence in them, at least (following my comparative presentation of Uniqueness) *vis-à-vis* other beliefs they entertain and that are not similarly disagreed upon. More specifically, reasonable agents need to separate the beliefs they hold that are not affected by reasonable disagreement from the ones that are so affected, and assign higher credibility to the first and lower to the latter. Which in turn implies, however, that reasonable agents must assign a suboptimal level of credibility to some of the values that, at least following the

¹⁸ See the distinction between ‘reasonable pluralism’ and ‘pluralism as such’ in PL, pp. 64-65.

¹⁹ On this position in particular, I agree with (Enoch, 2017, p. 140).

usual interpretation of comprehensive doctrines as conceptions of the good, provide the deepest source of meaning for their life.

This result is, as I mentioned before, paradoxical for political liberalism. Political liberalism originated out of Rawls's discomfort regarding the public justifiability of his model of justice as fairness. The challenge was to prove that a particular reasonable conception of justice, such as justice as fairness, could 'not only provide a shared public basis for the justification of political and social institutions but also helps ensure stability from one generation to the next' (Rawls, 1999 [1987], p. 421). The proposal was to rely on a stable consensus of reasonable conceptions of the good, which overlap on some 'political' values precisely because these values are independent of commitments to any conceptions of the good. In Rawls's own words, an 'overlapping consensus appears far more stable than one founded on views that express scepticism and indifference to religious, philosophical, and moral values' (Rawls, 1999 [1985], p. 413).

Our conclusion was that reasonableness demands a certain scepticism (although not of a full-blown nature) precisely towards the composing values of one's conception of the good. The consensus only comes about after people have started subjecting their conceptions of the good to reasonable doubts. Hence, the paradox of modelling public justification on the assent of citizens who already adhere to conceptions of the good, whilst demanding that they substantially revise their confidence in said conceptions. The paradox is relevant not just for being a stain of imperfection in the otherwise impeccable Rawlsian theory, but because it undermines both the idea of legitimacy prevalent within political liberalism and (as I will show in the chapter dedicated to that) the connection between this idea and the attainment of political stability.

On legitimacy itself, we can say that the paradox of reasonableness and scepticism provides further evidence of the implausibility of a model of legitimacy based on public justification. Again, the conclusions are similar to the ones of the previous chapter; adopting a standard of legitimacy that pays respect to the fact of pluralism leads to problems that political liberalism seems unfit to solve.

3.2 Conclusions on Legitimacy and Disagreement

The discussion of legitimacy and disagreement has concluded in a sceptical tone. The central question I have tried to answer in this and the previous chapters is whether a standard of political legitimacy can successfully overcome the problem posed by the existence of value-disagreement in society. That was, so to speak, the promise of political liberalism.

What I have shown is that all variants of political liberalism that pursue the grand project of reconciliation between disagreement and legitimacy are trapped in a dilemma. If they try to capture the real phenomenon of disagreement with attention to its nuances, they have to abandon some of the assumptions that are necessary to draw the most relevant normative conclusions. Once these assumptions are relaxed, however, the normative implications of 'public justifications' appear either impossible or at any rate extremely hard to be met. If, on the other hand, they retain such assumptions, even in the face of their empirical implausibility, then 'public justification' becomes no other than a general, 'axiomatic' constraint on political legitimacy.

I showed that these uncomfortable results obtain regardless of which model of public justification is adopted. Conditioning public justification on the non-dissent of citizens who can advance 'intelligible', sincere objections to a particular legislative

enactment generates a dilemma of its own. If we understand the proviso of 'intelligibility' rigidly, judging whether views advanced by others are intelligible becomes an overly demanding epistemic requirement; if we ignore it, however, public justification becomes virtually impossible as every conception of the good is potentially provided with a veto against the implementation of public norms.

Similarly, conditioning public justification on the consensus of reasonable agents can mean at least two different, and incompatible, things, depending on the epistemic demandingness of reasonableness. If reasonableness is a sufficiently robust epistemic requirement, then it will incorporate some form of scepticism regarding beliefs that are affected by reasonable pluralism, i.e. those beliefs that form part of one's conception of the good. On the other hand, if the definition of what counts as a reasonable consensus is substantially left to citizens' interpretations, that will generate an infinite chain of interpretive disagreements which will compromise great parts of the benefits of consensus, in terms of both political stability and *de facto* legitimacy.

What, then, of legitimacy and disagreement? The present work is not centrally about political legitimacy *per se*, but, as the last three chapters have surveyed the literature on legitimacy from various perspectives, I feel some conclusions are inevitable.

A standard of political legitimacy can be more or less independent from citizens' actual perceptions and ideas. Standards that completely abstract from citizens' ideas are theoretically valid, but ignoring citizens' actual consent risks turning legitimacy into a different notion such as justice. On the other hand, letting legitimacy be completely determined by people's actual consent is a philosophically naïve move. Leaving aside the infinite problems related to the generation of people's preferences

in circumstances of injustice, people simply do not express anything minimally comparable to a consent for public authority.

Citizens can however express some appreciation or lack thereof of the policies currently implemented. That is a point that I believe a theory of legitimacy cannot neglect and is the reason behind my inclusion of a general assessment of a political society's pattern of compliance within a plausible standard of legitimacy. An exclusive reliance on these evaluations, however, would lead to a purely voluntarist model of legitimacy, and, by implication, to the grim conclusion that most existing states are illegitimate.

I believe this forces the theorist of legitimacy into an inescapably binary choice. On the one hand, she can bite the bullet on the 'grim conclusions'. She will then defend a standard of legitimacy based on the common, and possibly unanimous, perception, among citizens in society, that both the constitutional order and the 'rough sum' of the norms requiring compliance are acceptable. On the other, she can use some objective standards to filter or reinterpret citizens' preferences and perceptions. The standard of legitimacy in this case would hold that a legitimate state is one in which citizens share a general perception about the acceptability of the state itself, *and* that it is also broadly acceptable according to objective, perception-independent criteria. This second model has a lot to explain in terms especially of how to differently weigh the separate objective and subjective components of legitimacy; it also has, though, a lot to offer,²⁰ especially in helping to determine which states in the world are more legitimate than others.

²⁰ See, for a recent attempt at deriving an account of political legitimacy from the assessment that citizens express towards their government, reasonably calibrated to discount cases of false consciousness, (Greene, 2016) and (Greene, 2017). But it is also worth pointing out the recently growing interest in the question about how distortions in the 'basic structure' of a certain society, understood in the widest possible way, and therefore encompassing more than its constitutional principles, can

I believe the analysis conducted in the last two chapters does not commit to any specific choice in this debate. It does commit, however, to the conclusion that one choice in this regard must be made and that philosophers working on legitimacy must choose whether the shared perception that a particular constitutional order is acceptable may entitle a state to the right to demand political obligations or whether more need to be said about those particular perceptions and the context in which they are produced. Both responses reject, however, the idea that public justification, both in the sense of consensus of reasonable doctrines and of convergence of intelligible ones, has any role in legitimacy, which is the crucial conclusion of this and the previous chapters.

determine a shift in the political obligations that some citizens are required to discharge as in (Shelby, 2007).

4. Equal Respect for Differences: An Oxymoron?

As I anticipated in the Introduction, two general principles will guide the positive response to the question of disagreement: equal respect and stability. It is now time to flesh both out, starting with equal respect.

The ideal of equal respect constitutes one of the central tenets of liberalism. However, there is little agreement among liberals about either the scope or the ground of the ideal; about what has to be respected equally, and why. Another form of consensus, albeit incapable of solving any substantive disagreement, concerns the historical origins of the ideal, which almost all scholars reconnect to Kant's practical philosophy. The Kantian background of the ideal, however, introduces problems of its own, not only because of the complex exegetical difficulties but also for the metaphysical assumptions that underlie Kant's practical philosophy in general, starting from the characterisation of moral agents as the elusive noumenal selves.

In the present chapter, my aim is luckily modest as I only have to address a small portion of the debate on equal respect. I wish to defend an ideal of equal respect that is appropriate to what I called in this work circumstances of endemic and irreducible value-disagreement. What kind of respect, if any, do we owe people in their role of public supporters of diverging conceptions of the good? An easy but ultimately impracticable response would be that, as we respect people equally, we ought to respect all of the values that they support and advance publicly; in the jargon of political liberalism, we can say we owe equal respect to all conceptions of the good. But equal respect for conceptions of the good cannot be unlimited. At the very least, we might want to preserve the right of a public authority to prevent the spread of views that are detrimental to the maintenance of the public order or that are morally repugnant.

In the rest of the chapter, I am going to defend a minimalist account of what we can call 'equal respect for differences'. I will rebut the scepticism expressed by some authors about this idea and limit the normative recommendations of equal respect to a *positive disposition towards an unbiased comprehension of others' views*. What we owe others in circumstances of public deliberation, I will argue, is no more than a willingness to try to understand their point, however mistaken or obnoxious it might appear to us. We owe this minimal form of respect to our fellows, especially those who share our political institutions, because we assume that they, like us, possess a capacity to form and revise a conception of the good. In other words, we assume that they, like us, are *autonomous, rational* persons that are capable of engaging with value in a way that is remarkably similar to ours.

I will present my account mainly against two alternative viewpoints. On the one hand, Kant's original view, remoulded in the latest years by scholars such as Alan Gewirth, grounds equal respect in a property equally possessed by all human beings, variously defined as autonomy or deliberative rationality. On the other, critics of the Kantian tradition, whether anti-liberal identity theorists or multiculturalist liberals, have turned equal respect into a principle about the equal prosperity of multiple communities of value co-existing under the same political institutions.

I will show why neither of the views succeeds. The Kantian view is incapable of producing an account that is not grounded in deep metaphysical assumptions. Those assumptions suffer from two inescapable flaws; they are generally hard if ever possible to justify, and they contrast with the empirical reality of human reasoning. Despite its flaws, I will take an important suggestion from the Kantian view, which is that, to respect people equally, we have to attribute to them a property that they all equally possess. But, I will argue, we do not need to attribute to people *actual*

qualities, of the kind that we can empirically assess. On the contrary, a purely hypothetical attribution will suffice, and that hypothetical attribution will constitute the core of my proposal. I will argue that what we need to attribute to others in circumstances where equal respect is required is a form of minimal rationality, that I will call *as-if rationality*. This is equivalent to a capacity to form and revise one's conception of the good or to claim authorship for the values one publicly endorses.

With this clarification in mind, I will proceed to dismiss the views recommending unmediated respect for conceptions of the good. I will reject in particular the view according to which equal respect implies a formal recognition, on the part of the public authority, of the different conceptions of the good existing in society. The failure of this model lies in its demand that the state engage in an active policy of evaluations about citizens' conceptions of the good.

As I will show in the dedicated section, in contrast with what some of its advocates argue, unmediated respect is a scarce resource whose attribution to some communities always implies deprivation to others. Hence, the public authority needs to engage in evaluations in order to determine which of the various communities of value in society are entitled to this form of protection and can therefore prosper. This policy of evaluations contrasts with the normative requirement, which I derive from my model of 'as-if rationality' respect, to abstain from epistemic evaluations about people's possession and use of their capacity to form and revise a conception of the good.

I will conclude by noticing how, despite its minimalist nature, respect as attribution of 'as-if rationality' is still capable of demanding a specific *attitude*. This corresponds, as mentioned above, to the willingness to make sense of others' views in the most unbiased possibly way. Despite being mainly an attitudinal requirement (the full comprehension of others' views cannot be directly required), my model of

equal respect does give rise to both positive and negative moral duties, which I will present in the final section.

4.1 From Kant's Account to a Kantian Ideal

4.1.1 Kant's Original Account and Its Difficulties

In this first section, I am going to start by presenting Kant's account of equal respect. I will show that the account does not furnish satisfactory conceptual resources to organise a plausible response to the present investigation. However, following suggestions coming from Rawls and other Kantian contractualists, I will build a 'Kantian' ideal of equal respect that, though deprived of some problematic metaphysical and epistemological assumptions, aims at developing Kant's own premises.

The most articulated presentation of respect (*Achtung*) in Kant's work is in the *Metaphysic of Morals*, where Kant grounds respect in human dignity, a notion that had already played a prominent role in the prior *Groundwork of the Metaphysics of Morals*. In the *Groundwork*, Kant had defined dignity as the worth specific of moral agents ('persons'), who are characterised by their being capable of rational deliberation; this capacity allows to consider persons as *ends in themselves*, rather than exchangeable, priced means. In the *Groundwork*, however, he had merely employed the term 'respect' to define the feeling moral agents ought to express with regard to the moral law.¹

In the *Metaphysics of Morals*, on the other hand, Kant firstly expounds the threefold connection between respect, dignity and personhood (the normative notion of being an end in itself) in a passage of impressive clarity;² he then classifies

¹ See in particular (Kant, 1996 [1785], p. 56) for Kant's description of 'respect of the law'. See (Broadie & Pybus, 1975) for further reflection on the recipients of respect in Kant's moral philosophy.

² '[A] human being regarded as a *person*, that is, as the subject of a morally practical reason, is exalted above any price; for as a person (*homo noumenon*) he is not to be valued merely as a means to the

respect as one of the two maxims correlated to the expression of duties to others (the other being love). Respect is defined as the 'maxim of limiting our self-esteem by the dignity of humanity in another person' (Kant, 1996 [1797], p. 569), and, more importantly for our analysis, is distinguished from love by its being due to every person, regardless of their moral (as opposed to human) worth. Kant seems to imply that what is properly respected in humanity is the potentiality of rational deliberation, equally possessed by every human being, rather than the use of that potentiality on the part of individual human beings. It is the possession of the capacity itself that grounds the practical maxim to treat every person not merely as a means to someone's ends, but also as an end in itself.

This formulation of respect for persons may seem to give rise exclusively to negative duties, namely the negative duty to refrain from considering oneself and other persons as mere tools in the achievement of specific goals (such as the pursuit of one's or others' wellbeing).³ In a significant passage, however, Kant provides an interesting casuistry of the way through which individuals who have failed to use their capacity of rational deliberation successfully can still be respected.

On this is based a duty to respect a human being even in the logical use of his reason, a duty not to censure his errors by calling them absurdities, poor judgement and so forth, but rather to suppose that his judgement must yet contain some truth and to seek this out, uncovering, at the same time, the deceptive illusion (the subjective ground that determined his judgement, which, by an oversight, he took for objective) explaining to him the possibility of his having erred, to preserve his respect for his own understanding.

ends of others, or even to his own ends, but as end in itself, that is, he possesses a *dignity* (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world.' (Kant, 1996 [1797], p. 557).

³This is also Kant's claim (Kant, 1996 [1797], pp. 569,583).

For if [...] one denies any understanding to someone who opposes one in a certain judgement, how does one want to bring him to understand that he has erred? (Kant, 1996 [1797], p. 580)

What transpires from this quotation is that respect for persons entails a duty to interpret others' beliefs *as if* they were the product, however mistaken, of rational deliberation. Alongside this positive aspect (the willingness to interpret others' beliefs as consistent with a path of rational deliberation that has failed at some point), there are negative implications surrounding the duty, such as the requirement not to consider the mistaken view as an 'absurdity', a belief that lies outside the realm of rationality.

If taken to represent a direct response to the present investigation, this view seems implausible. It implies that all human beings, purely because of their personhood, possess a capacity for rational deliberation that they are more or less successfully employing when they advance a (moral) 'judgement'. Identifying this ubiquitous capacity however looks problematic. Kant's attribution of rational deliberation to all human beings is deeply grounded in what is arguably the most metaphysically demanding component of his moral theory, namely the dualism between a fallible and limited theoretical reason and an infallible and universally accessible practical reason. According to Kant, people behave morally when they let their conduct be determined by practical reason, so that the moral law will eventually condition their behaviour in the same way that the laws of nature condition the world outside human brain (the difference being that, unlike animals, humans let their conduct be determined by a law they have freely assented to). Practical rationality therefore equates, for Kant, to full adherence to the moral law.

But, regardless of the infinite complexities in justifying the belief in the infallibility, on the one hand, and complete intelligibility, on the other, of practical reason,

this attribution of practical rationality is irrelevant to the present discussion. People in circumstances of endemic and irreducible disagreement are not behaving in a way that suggests they are all trying to respect the same moral law; on the contrary, they are advancing the most various moral judgements, most of which are mutually contradictory. Hence, either we come to the conclusion that they are all irrational, or we have to accept that, if we want to take the fact of pluralism seriously, Kant's notion of practical rationality must be adequately amended.

4.1.2 The Problem of the Identification of the Properties on Which Respect Might Supervene

Numerous authors in the last decades have tried to separate Kant's account from its metaphysical grounds. They have replaced the Kantian category of rational deliberation, with its contentious metaphysical bases, with others that are presumed to be less metaphysically laden. Two argumentative routes seem available. One consists of identifying an empirical or quasi-empirical (as opposed to metaphysical) property that is indeed possessed by (roughly) all human beings, at least at a certain point in their ordinary cognitive development, and to use this, rather than the Kantian rational deliberation, as the ground of equal respect. The second, which I will follow, consists of sidestepping the identification of an empirically shared property and ground equal respect instead in a property that is merely attributed to, without being necessarily possessed by, all human beings (again, at least starting from a certain moment in their life). I believe this second strategy is implicit in the most successful contemporary adaptations of the Kantian ideal.

Before I come to that, however, I need to explain why I believe the first route, identifying a quasi-empirical property that is possessed by all human beings, is hopeless. One reason why this might be so is the objective difficulty in identifying

such property. I do not take this, however, to take an insurmountable difficulty. In order to have at least some empirical feature, the property must be something whose possession can be determined via an empirical investigation, albeit of a very unsophisticated kind. For example, we might include properties that, despite not being empirically traceable (we do not know whether they have a physiological correlate), are still recognisably possessed by all or most human beings, in the sense that their possession is necessary for having certain experiences that (almost) all human beings share.

By following this procedure, we arrive at fundamental capacities that characterise the ordinary agency of human beings and without which agency itself is unexplainable. Alan Gewirth suggested the notion of 'purposiveness' in this regard, that is, the ability to act following purposes one has freely assigned oneself. For Gewirth, human dignity is grounded in the quasi-empirical observation that human beings, in order to act, must be capable of valuing their own ends; in so doing, they must also attribute worth to themselves as the source of ends that motivate their action. Moreover, every moral agent must also attribute the very same worth to every other 'actual or prospective agent' '[f]or their actions have the same general kind of purposiveness that provides the ground for his attribution of dignity to himself' (Gewirth, 1992, p. 23). The 'generic conditions' that make human agency possible, therefore, constitute for Gewirth the equally distributed property that can ground the normative attribution of dignity to all human beings.

Are the 'generic conditions' of agency a plausible ground for the universal attribution, across the human species, of dignity and respect? There are many objections to it. Firstly, it is worth mentioning how the belief in human beings' capacity to assign themselves purposes that are not pre-determined is in itself a metaphysical assumption, and, incidentally, one of the most controversial. One could avoid this

problem by simply claiming that what matters is the perception of acting following self-assigned purposes and not the underlying metaphysical explanation of the phenomenon; even if acting following self-assigned purposes was a huge delusion, it would still be one shared by (almost) all human beings, and that is what matters.⁴ But, if it is the case, then we have to exclude from equal respect all those members of the human species who do not entertain that perception, maybe due to a cognitive handicap or trauma, and include instead a cyborg that has been programmed to feel that way, despite the fact that all its actions and perceptions are pre-determined.

More importantly, however, a proposal like Gewirth's falls prey to an arbitrariness objection. Even if we could indeed stipulate that all humans, at least at some point in their cognitive development, are creatures that inherently value their ends and must therefore ('on pain of contradiction') value their ability to set ends, why would we want to use precisely that feature of human behaviour as the basis for the attribution of equal respect? Presumably, human agents share a plurality of properties, many of which have been traditionally characterised as definitive of the human species, and practically all of which are progressively developed in human life through both education and growth. Take, for example, such things as the use of language as a vehicle of communication, the ability to produce reasons that others can engage with and respond to, the capacity to create relational bonds that overcome simple natural needs, and many others.

An immediate response to this kind of arbitrariness objection would be that purposiveness is more primitive than any of the other properties because the expression of every other human ability must depend on the ability of an agent to set oneself ends. Indeed, Gewirth affirms that 'since the agent unforcedly chooses to do

⁴ A commentator observes that, for Gewirth, the valuation of purposiveness 'presupposes that agents at least *feel* that *they* choose their purposes, whether or not they actually have freewill' (Beyleveld, 2014, p. 232). (Emphasis in the text.)

x for the purpose of attaining or having E, he must think that his purpose E has sufficient *value* to *merit* his moving from quiescence to action in order to attain it' (Gewirth, 1992, p. 21 – Emphasis in the text.). This response, however, turns purposiveness into a trivial notion, and one that cannot even characterise the human species as such. Even a cat, when she moves from quiescence to action in order to attain the simplest end (for example, getting food) must have an extremely preliminary notion that the end she is trying to attain is good enough to motivate her to act. Of course, one could counter, following Kant strictly in this regard,⁵ that the difference between the human and the cat lies in the fact that the cat hasn't set herself any end, but has derived an end from her nature. But this seems no other than a metaphysical presupposition; what's worse, a metaphysical presupposition that sharply clashes with anything that evolutionary biology and psychology have produced since Darwin's time. Hence, a *prima facie* plausible response to the arbitrariness objection has triggered an objection that is much more detrimental, as on the one hand it trivialises the property on which equal respect should supervene and on the other risks compromising the 'human' limitation of equal respect.

A further problem affecting this kind of paradigm is that human abilities (of any kind) are possessed unequally. Different human agents, even after they have reached the age supposedly necessary for full cognitive development, possess them at different degrees. This point about the scalar possession of abilities might not af-

⁵ Kant had indeed suggested that the '[t]he power to set an end – any end whatsoever – is the characteristic of humanity' (Kant, 1996 [1797], p. 392). The reason for why Kant uses the adverb 'whatsoever', and seems therefore to infer that even setting ends that are contrary to the moral law is a feature of humanity, is because he takes as the primary character of humanity the ability to overcome instinctual inclination that derive from nature. In this sense, humanity differs from 'personality', which is not the capacity to set whatsoever ends, but to set ends that are derived from practical reason and are consistent with the moral law. For this reconstruction, see (Korsgaard, 1986, p. 189).

fect the most trivial sense of purposiveness, because we can well suppose that virtually all humans (bar the ones that, for example due to their comatose state, are prevented from autonomously acting in any way) can set themselves (however minimal) purposes. But it will definitely affect all non-trivial human abilities on whose possession someone might want to ground equal respect. Does this entail that equal respect must then be attributed to human agents *in proportion* to their possession of the quasi-empirical property on which respect supervenes? If that is the case, the very notion of an *equal* respect seems lost for good.

The supervenience of equal respect on particular human abilities tends to generate a dilemma. Either the ability is so minimal as to be shared beyond the human species, or it is an ability that is distributed unequally even among (adult) human beings, which seems to entail that respect, *contra* the original Kantian model, is not assigned equally and universally to every human being. The first interpretation gives rise to problems of over-inclusiveness as it is incapable of circumscribing the attribution of respect to *human* agents, the second, to under-inclusiveness at least in the sense that some human beings will be excluded from the attribution of maximal respect.

Somebody who wanted to solve the problem of over- or under-inclusiveness for good might as well take the *Homo sapiens sapiens* genome as the equally possessed property grounding respect. Is it less arbitrary than, say, the (almost) equal sharing of the perception of purposiveness? Maybe so, as one could always claim that the perception of purposiveness is necessary for agency, whereas the use of the *Homo sapiens sapiens* genome to ground respect would be a form of utter speciesism. But, is there a non-speciesist way to limit the problem of over-inclusiveness? Notice that my problem is not with the possibility of respecting non-human animals, but with elaborating a form of respect that can be specifically applied to a particular type

of typically *human* interaction, namely deliberation – through values – in the public sphere.

4.1.3 A Shift in Paradigm: Equal Respect without Empirical (or Metaphysical) Foundations

I propose to change the paradigm. We do not need to rely on either arbitrary empirical features or unattractive metaphysics to ground an ideal of equal respect for differences. Instead of interpreting the Kantian account as a theory about the bases of equal respect, we see it as an account having a centrally *normative* function, which is to define certain modes of treatment of others as consistent or inconsistent with the respect that is due to humanity as such.⁶

The three core components of a Kantian ideal of equal respect that we can isolate as undertaking this core normative function are *universality, as-if-rationality, and independence from evaluation*. I will now flesh out each in turn.

Universality concerns the application domain of the ideal, or, in other words, the recipients of respect. In Kant, the individuals who are owed equal respect are all those capable of moral deliberation; this is supposed to encompass the human species as a whole⁷ (as well as other rational beings, were they existent). As we have seen, however, the deployment of rationality as the feature definitive of humanity

⁶ I can hardly claim to be the first person who suggests a similar shift. In the following paragraphs I am going to provide numerous references to scholars that have inspired my reassessment of the Kantian account. Other than them, another author who has attempted a task altogether similar to mine, although with the aim of defending an account of equal respect designed to ground the equal attribution of democratic rights, is Valeria Ottonelli (Ottonelli, 2012). See in particular this sentence: 'I suggest that, if we want to overcome these problems, we must disentangle the idea that equal respect can serve as a principle of democracy from any factual assumptions about citizens' equal competence. This entails shifting the phrasing of the principle. Equal respect should be interpreted as a requirement of respectful treatment, rather than as a requirement of sincere responsiveness to competence' (p. 211).

⁷ 'Every human being has a legitimate claim to respect from his fellow human beings and is in turn bound to respect every other' (Kant 1996 [1797], p. 579).

as a species is based in Kant on a metaphysical premise that we would prefer to avoid.

Hence, if we want to maintain universality, we have to give up on any attempt to define the ‘bases of equal respect’⁸ in any empirical or metaphysical way. We still need, however, to attribute to others a characteristic feature that might explain our equal treatment in a way that is not arbitrary or *ad hoc* (which was the problem I identified before in the attempts to provide purely empirical bases to equal respect). We need therefore to assume that a certain form of rationality, which I will later specify, is indeed shared universally among persons. I call this attribution *as-if rationality*, as opposed to ‘rationality potential’, to underline further its non-empirical character; even assuming that every human being is potentially capable of rational deliberation would be empirically false. As-if rationality, in other words, bypasses the insidious problem of trying to identify a capacity or capability whose possession triggers the attribution of equal respect, and it does so without relying on an empirical notion of potentiality which may re-introduce the same problems it is supposed to overcome.

I call the final element that characterises the Kantian ideal ‘independence from evaluation’ and we can easily see it as logically implied by the combination of the former two. If equal respect has to be accorded to every member of the human species merely because of an assumption of rationality, it follows that the attribution

⁸I derive this phrase from (Carter, 2011).

of respect must be independent of evaluations of single individuals' moral or cognitive worth.⁹ In fact, the only worth that matters in Kant's account is *dignitas*,¹⁰ the worth characteristic of humanity *per se*. Persons possess *dignitas*, for Kant, in virtue of their capacity for rational deliberation. As in the Kantian account I am defending rational deliberation is replaced by 'as-if rationality', the bases of human dignity are also to be found in 'as-if rationality'. In both accounts, the variations in the actual use of rationality, whether they lead someone to moral rectitude and ingenuity, or whether they lead to moral evil or stupidity, do not affect the entitlement to equal respect.

I believe the argumentative strategy I have employed is implicit in some of the most promising accounts of equal respect, starting from Stephen Darwall's. Darwall famously distinguished forms of respect based on 'recognition', of which equal respect for persons is an instance, and forms of respect based on 'appraisal', which emanate from an evaluation about the way in which an entity manifests excellence of some kind. The distinction is linguistically misleading because appraisal implies a form of recognition ('I recognise that you have excelled in x, and that entitles you to appraisal-respect'). But the two forms of recognition crucially differ. In its being a recognition of excellence, 'appraisal' is premised on a prior positive assessment of desert, whereas 'recognition' proper (in Darwall's sense) corresponds

⁹ I therefore disagree with Axel Honneth's reconstruction of the Kantian conception of 'recognition', which he reconnects to an 'evaluative perception, a capacity for which every adult who has been socialized successfully normally has at her disposal' (Honneth, 2001, p. 125). Honneth seems to imply, contrary to my suggestion before, that according respect amounts to a form of perceiving worth in another person, rather than merely according certain properties.

¹⁰ In German, the words 'worth' and 'dignity' both translate with the same term, 'Würd'. That is why Kant uses the Latin 'dignitas' to create a distinction.

to the unmediated attribution of *status* to a specific object.¹¹ Darwall, following insights from Joel Feinberg and P.F. Strawson,¹² sees respect as a second-personal attitude that aims at recognising in other a specific authority or standing, namely the authority to make valid claims.

This recognition corresponds to a form of *attribution* to others of a particular normative property, that of being a person, as Darwall explicitly admits in writing that ‘we respect someone as a person when we accord her this second-personal authority’ (Darwall, 2004, p. 51). The attribution of authority presupposes, however, that others are *capable* of exercising authority;¹³ that they are capable, in other words, of producing intelligible claims against which I can be held accountable to. This presupposition, which is implicit in Darwall, is what I mean by ‘as-if-rationality’.

‘As-if rationality’ is, in the view I am defending, the minimal cognitive feature we have to presuppose in others in order to ground our respect for them; absent that presupposition, we would not have any reason to let their claims condition our public conduct and deliberation. The type of rationality we attribute must not be too

¹¹ Darwall did not make this explicit in his original presentation in (Darwall, 1977), where he simply limited himself to the description of ‘recognition respect’ as ‘a disposition to weigh appropriately in one’s deliberations some feature of the thing [to be respected] and to act accordingly’ (p. 38). In the more recent (Darwall, 2004), however, the point about ‘recognition respect’ as primarily a recognition of status becomes explicit.

¹² See especially (Feinberg, 1970) and (Strawson, 1974).

¹³ Somebody might disagree. In a recent intervention, Laura Valentini seems to hold a view according to which, when discussing the foundations of human rights, the normative and descriptive components of human dignity (which she refers to as ‘status’ and ‘inherent dignity’ respectively) can be kept separate: ‘I argue that the focus should shift to status dignity—particularly, the status dignity of individuals *vis-à-vis* powerful sovereign entities—*leaving open what inherent human property grounds or justifies that status*’ (Valentini, 2017, p. 863). I believe, in contrast to this attempt at keeping the two components separate, that the attribution of respect must be premised in a certain understanding of others as somehow *deserving* (in an extremely minimal sense) that respect. Absent that premise, respect for others would turn from an attitude involving trust (see footnote 16) to a leap of faith involving completely *blind* trust.

demanding, lest we attribute to fellow human beings a feature that too many of them do not possess or exhibit. As the present investigation broadly concerns the treatment of people in circumstances of public debate or deliberation, we only need to attribute what is minimally necessary for people to participate in those circumstances. What is minimally necessary is that people are capable of advancing their own conception of, to summarise crudely, both *what matters* to them and *what is*, in their view, *right or wrong*. To echo a Rawlsian formulation, we conclude that, in order to respect people ‘as if’ they were capable of rational deliberation, we take them as capable of forming and revising both a conception of the good (on which I will focus in particular in this chapter) and a sense of justice.¹⁴ In attributing our fellow human beings ‘as-if rationality’, we attribute to them the two Rawlsian moral powers.

In sum, respect in the Kantian model I am defending is both *expressed through* and *grounded in* the attribution to our fellow human beings of the faculty of ‘as-if rationality’. The next sections of the chapter will try to derive from these vague and general principles some more direct normative recommendations.

A final explanation is necessary. What I still need to explain is the reason why attributing to everyone a minimal form of rationality is preferable to taking into account the different expressions of that capacity, which would lead to accepting variations in the expressions of respect. To respond to this issue, I now intend to analyse my attribution of ‘as-if rationality’ to Ian Carter’s notion of ‘opacity respect’ (Carter, 2011; Carter, 2013).¹⁵ As in Carter, our account indeed mandates that, in

¹⁴ See (Rawls, 2005 [1993], pp. 29-35) for the ‘political conception’ of persons possessing the two moral powers.

¹⁵ Notice that my ‘independence from evaluation’ parallels in this respect Carter’s notion of ‘evaluative abstinence’ (p. 550).

order to treat someone with respect, we refrain from analysing those cognitive aspects of one's character that overcome a minimal threshold of what Carter calls, following Rawls, 'moral personality'. As moral personality is, for Carter, a range property (a binary property that one either possesses or does not) that in turn supervenes on the possession of rationality (a scalar property) over a certain threshold, the differences in the exercise of rationality above the threshold of moral personality are morally irrelevant. Not only that, respect for people's dignity demands, in Carter, that such investigations in the level of rationality be morally prohibited; dignity implies treating others as 'opaque' by 'adopting a perspective that avoids evaluation of the agential capacities on which moral personality supervenes' (Carter, 2011, p. 552).

'As-if rationality' is, in a sense, a development of 'opacity', and, in another, a departure from it. It is a development of 'opacity' because, consistently with the normative requirement of opacity, it pays respect to people's dignity by imposing a strict limit on the morally acceptable investigations one can undertake to determine others' level of rationality. But it departs from Carter's original model insofar as it accompanies the negative component of restraint with the positive demand to attribute to others a property they do not necessarily possess (empirically speaking), which is rationality. Attributing 'as-if rationality' amounts to acknowledging that respecting others equally includes an element of trust;¹⁶ our respecting them is not based on our *knowing* that they possess features that render them the appropriate object of our respect, but on our assuming that at least one such feature (a minimal form of rationality) is indeed shared by them.

Once the conceptual grounds of equal respect have been clarified, my aim for the next two sections is the following. I will first (section 2) demonstrate why two

¹⁶ On the relationship between trust and equal respect, see especially (Wolff, 1998, pp. 108-109).

different attempts to flesh out an ideal of respect for differences that have been presented in the literature fail due to their violating some of the principles I identified as central to the Kantian account. Then (section 3), I will provide a detailed presentation of how my account of equal respect as attribution of ‘as-if rationality’ plays out in the circumstances of value-disagreement that are the focus of the present work.

4.2 Respect in Circumstances of Disagreement: Substantive Accounts

4.2.1 ‘Justificatory Liberalism’ and the Principle of Sincerity

My analysis of the literature begins with the tradition of ‘justificatory liberalism’ I have already copiously cited and its ‘principle of sincerity’. At the end of the section dedicated to it in chapter 2 (§2.1), I suggested that maybe one way to rescue the entire project of ‘justificatory liberalism’ if we interpret it as an attempt to produce public norms that are an expression of public respect for differences. I will start the analysis with Gaus’s ‘principle of sincerity’:

Betty’s argument justifying *N* to Alf is sincere if and only if (1) she is justified in accepting *N*; (2) she has a justified belief that *N* is justifiable in Alf’s system of reasons and beliefs. (Gaus, 1996, p. 141)

The principle straightforwardly applies to circumstances of endemic disagreement and it does include a clear normative recommendation. It mandates taking into account the perspective of others and what may be, from their own viewpoints, acceptable. Specifically, if a certain argument is not justifiable from *x*’s subjective per-

spective, sufficiently adjusted to rule out internal inconsistencies and idiosyncrasies,¹⁷ and x is a member of practice P, then the argument, whatever its content and structure, can never aspire to public justification within P.

In order to turn 'sincerity' into a full-blown principle of equal respect, we can follow some suggestions coming from Kevin Vallier,¹⁸ and imagine establishing a system of entitlements that aim at protecting citizens who can advance 'intelligible' complaints against the implementation of coercive norms. The system will imply that, if citizen x can advance a claim against norm N that is 'intelligible' to the citizenship at large (in the sense that everyone can see the relevance of the claim for x's subjective belief set), then the norm will be rejected.

We can further claim, here slightly beyond what Vallier explicitly says, that this system of entitlements is the proper way to express equal respect in circumstances of public deliberation. An ideal of equal respect leading to these normative recommendations is consistent with the Kantian account. The main normative recommendation we derived from the Kantian account was to treat people as if they were capable of forming and revising a conception of the good and a sense of justice. It might now seem that a plausible way to ensure that people are respected in these roles is to implement a system of entitlements that ensures that subjective values are capable of influencing political deliberation. I call this the 'veto-conception' of equal respect as it practically consists in providing citizens with an equal veto towards the full implementation of norms that collide with the values that they, in their capacity of *as-if* rational deliberators, can advance.

Despite its consistency with the Kantian account, this apparently neat attempt at cashing out the ideal of equal respect has its shortcomings. Regardless of

¹⁷See §2.21.

¹⁸(Vallier, 2011) (Vallier, 2016 b) (Vallier, 2016 a).

the legal complexities, which I will consider in the final chapter, the 'veto-account' is implausible as a possible explication of the ideal of equal respect as it imposes duties that are virtually impossible for the involved agents to discharge.

As an example, take the debate on the moral and legal admissibility of abortion. If all the agents participating in the debate decide to adopt the principles of sincerity and justification through 'intelligible' non-rejectability¹⁹ as the ideal driving their conduct, the likeliest situation to come about is one in which the debate is blocked at the outset. Indeed, neither the pro-life nor the pro-choice supporters can easily advance arguments that they sincerely believe will be justifiable to their opponents' belief systems. The key lies in the concept of sincerity; the debaters may advance all sorts of arguments that they presume are *prima facie* appealing to their adversaries, but how often is this attitude the result of a mere misconception of the other's actual position, rather than an attentive and unbiased analysis?

Suppose I am a pro-choice activist and I advance the argument that, at least in case of danger for the mother's life, or in case of rape, abortion should be permitted. Suppose further that, in advancing this argument, I refrain from mentioning in its support any considerations derived from the general moral framework I submit to (such as one that prioritizes individual autonomy or self-ownership over the alleged value of human life). I say, instead, that everyone should be concerned about a mother's health conditions, and that in case of rape the fact that the pregnancy was involuntary certainly must have some bearing in the discussion.

There are at least two problems with my suggestion. The first concerns the extremely limited character of the compromise; if I am a 'pro-choice' supporter, then of course I do not want the state to deem abortion permissible only in cases of rape or risks on the mother's health. The second, more fundamental, problem concerns

¹⁹ I have obviously taken this label from Scanlon's test of 'reasonable non-rejectability'.

the *de facto* lack of sincerity in my proposals: if I had attentively analysed the 'pro-life' position then I would know that at least the rape argument is not appealing at all to someone who believes in both the sacred value of human life and the equality of the value of all lives from the moment of conception. Starting from those premises, it is easy to argue that what a mother can receive in case of rape is certainly sympathy, and help in some form, but never the right to take the life of a being who is as much entitled to the value of life as she is. The case of substantial threats to the mother's health is more complex, but it is at least possible to imagine that, in case one had to choose about which life to privilege, the mother's or the foetus's, the 'pro-life' supporter would not automatically accept saving the mother in any possible circumstance.

A system of rights that protected citizens' capacity to form 'intelligible' complaints against the law, based on subjective values, would then lead to a justificatory impasse. Some contingent features of abortion legislation, such as the fact that *any* kind of regulation could be opposed through potentially intelligible reasons and that allowing exemptions has here a limited practical impact, will imply that abortion legislation could never reach the status of public justification. Which in turn means that the 'veto conception' of equal respect commits one to the uncomfortable conclusion that abortion legislation (of any kind!) will inevitably deprive some people, those that can express intelligible complaints against the law, of equal respect.

Of course, this is not a decisive objection against the model. The defender of the 'veto conception' could still defend the view as the most plausible in the face of failing alternatives, and could further define the violation of equal respect in the abortion case as a mere *pro tanto* wrong, overridden by the necessity of abortion legislation. But the Kantian character of the proposal has got lost once the implications have become evident. The Kantian conception does attribute to every agent a

capacity for rational deliberation but does not assume that the capacity may not be misused. By contrast, the 'veto conception' attributes to citizens' beliefs by themselves, only minimally adjusted, the power to determine the rightness or wrongness of a scheme of public justification.

It is further noteworthy how, to respond to the objection about the proliferation of justificatory impasses, 'justificatory liberals' usually emphasise the role of the state as an 'umpire' or 'arbitrator' capable of breaking the indeterminacy caused by the impasse.²⁰ But, in case the adjudication is left to the discretion of the arbitrator, how could the agents who have been defeated in the deliberation receive equal respect?

The only justification offered to them would be that, due to the indeterminacy caused presumably by the intensity of the debate and the distance of the various positions advanced, the 'arbitrator' deliberated against them. Had the debate been less intense and the various positions more compatible, the 'arbitrator' would have been unnecessary and the discussion would have easily led to the enactment of publicly justified norms. But this result seems at odds with what we expect from an account of equal respect. We would expect that fellow human beings particularly deserve a demonstration of equal respect when they are engaging in a debate that deeply affects their conscience, whereas this account implies that those are precisely the scenarios in which, due to the impossibility of public justification, the 'veto conception' of equal respect is inapplicable and the recourse to the Leviathan-like 'arbitrator' is made necessary.

²⁰ (Gaus, 1999, p. 279); (Vallier, 2016 a, p. 612).

4.2.2 Respect as 'Hands-on Recognition'

Some authors have proposed instead to cash out an ideal of respect that is specifically designed to tackle the daunting problem of asymmetrical relations among different communities within a political society. The point of the ideal would be to transform the dynamics existing among the different components of society from inherently unfair or exploitative to respectful, so to create a society in which the differences among groups are maintained without giving rise to conflicts and injustices. The central idea is that differences deserve, within democratic societies, a positive recognition, a form of 'hands-on' respect, which their proponents see especially in contrast to difference-blind, classic liberalism.

Underlying this strand of thought is usually a global scepticism about liberal neutrality. Following Iris Marion Young's seminal work, we can identify three sources for this scepticism.²¹ Firstly, 'identity' theorists question the neutrality of 'difference-blind' standards of judgement and see these instead as a way to perpetuate privileges and stereotypes against members of oppressed minorities. Secondly, they question the neutrality of norms that are supposedly insensitive to cultural paradigms, and see these instead as concealed forms of 'cultural imperialism'. Thirdly, they question the effective benefit of neutrality towards individual members of oppressed groups, who, in their view, end up being doubly stigmatised, in their community for having renounced their cultural identity, and outside the community, where they are still seen as no other than members of disadvantaged or oppressed groups.

This three-pronged scepticism about liberal neutrality characterises the most radical attitude of authors that end up rejecting the liberal paradigm alto-

²¹ See (Young, 1990, pp. 101-105).

gether. Scepticism about cultural insensitivity has, however, vastly influenced liberal philosophy, leading to the development of culture-sensitive approaches within liberal philosophy.²² Charles Taylor, for instance, substantially agrees with Young in proclaiming the ‘procedural’ version of liberalism ‘guilty’ of being ‘inhospitable to differences’ with its insistence on ‘uniform application of rules’ and hostility to ‘collective goals’ (Taylor, 1994, p. 34).

At the normative level, both ‘identity’ critics of liberalism and culture-sensitive liberals share the belief in the inadequacy and limitedness of the system of traditional individual rights. Liberal theorists then propose to augment the traditional individual rights with a system of ‘cultural’ or ‘minority’ rights capable of guaranteeing protection not only to people’s abstract capacity to lead a life informed by values they have adhered to, but also to the ‘societal cultures’ (Kymlicka, 1995, p.

²² See in particular (Taylor, 1994) (Galston, 1995) (Kymlicka, 1995). Kymlicka has substantially buttressed his position, without significant revision, in (Kymlicka, 2001) and (Kymlicka, 2007), whereas recent defences of liberal multicultural rights are in (Galeotti, 2002) and (Patten, 2014).

I should point out that other liberal theorists, especially those defending so-called social or relational egalitarian views, have reached conclusions superficially similar to those of the difference-sensitive liberalism of Kymlicka and Taylor, insofar as they also advocate some forms of protection for groups. However, their concern is different; their critical focus is on to the damage of ‘social exclusion’, or, in the words of Jonathan Wolff, ‘differential social inclusion’ (Wolff, 2017) that affect members of marginalised or discriminated groups. ‘Groups’, however, is here understood simply to refer to a set of people united by the most diverse common feature, from the biological and psychological (women or LGBT people) to the social (the working class or the migrants) to the religious (the Muslim) to the ethno-cultural (the Roma or the French-Canadian). To put it in a slightly simplified form, the main preoccupation of social egalitarianism is not the protection of difference, but, by contrast, the achievement of unity, for which we mean ‘a society in which each person regards and treats each other as equal’ (ibid., p. 165). The main contribution in this strand of literature are (Wolff, 1998) (Anderson, 1999) (Anderson, 2010) (Hinton, 2001) (Hurley, 2003) (Scanlon, 2003 d) (Scheffler, 2003) (Satz, 2007).

I believe the conclusions of this chapter are broadly consistent with social egalitarianism, but I leave this demonstration to another work. What I feel I need to specify here, however, is that, in my critique of the cultural rights approach as a plausible expression of equal respect, I am not taking issue with *any* attempt to provide particular entitlements to people belonging to particular discriminated groups (understood in the most generic possible way), but only with the contention that particular cultural contexts (in Kymlicka’s words, particular ‘societal cultures’) *must* be the object of recognition and respect on the part of the public authority.

76) that provide their members with those very values and meanings. Notice that by a 'societal culture' we do not mean here something as abstract as a conception of the good that people from different backgrounds can freely adhere to, but a concrete organisation formed of shared practices and values that (a) develops and flourishes in a specific territorial and historical context and (b) presents more or less strict limits on its membership.

Kymlicka's argument is particularly interesting, as it does not depend on any form of valuation of cultural diversity *per se*. It depends, on the contrary, on the recognition, usually considered unproblematic in liberal theory, that 'societal cultures' have a fundamental role in giving people a familiar context through which they can exercise their freedom. That context might include things such as a common language, a possibly plural but finite set of values people make use of when they make 'significant' choices in their life, and more generally a system of cultural and social bonds capable of giving its members 'identification'.²³

Provided the possibility of revising and even rejecting that context is guaranteed, liberalism is certainly not incompatible with the acknowledgement that the forced separation of a person from her familiar context represents one of the most devastating harms human beings can suffer. However, most classic liberals would argue that the same rights and provisions that guarantee the freedom to *change* one's familiar context are also the ones that guarantee the freedom to *maintain* one. Those are usually identified with individual rights, and in particular complete freedom of expression, and, although this second point is less straightforward, some

²³ Kymlicka mentions (Margalit & Raz, 1990) as being an inspiration for his belief in the relevance of 'familiar context'. However, the thrust of Margalit and Raz's argument is different as they try to provide a justification for national self-determination based on the (controversial) identification of nations with those 'encompassing groups' that provide people with their familiar context. The role of the 'familiar contexts' in liberal political philosophy is further evidenced in (Raz, 1995 a).

form of warranty, on the part of the political authority, that citizens are exposed, mainly through education, to a plurality of 'familiar contexts'.²⁴

Kymlicka introduces, against the classic liberal paradigm, an asymmetry between those rights and provisions that are necessary for the *revision* and those that are necessary for the *maintenance* of the familiar contexts and uses that asymmetry to justify, from a liberal perspective, the recourse to 'minority rights'. The distinction between the culture-sensitive approach advocated by Kymlicka and the classic liberal approach can be so summarised. Classic liberals, despite their frequent caricature as defenders of absolute individual atomism, do not neglect the relevance of familiar contexts for individuals; they can well acknowledge the fundamental role familiar contexts play both in enhancing people's wellbeing and in allowing them to exercise their moral powers fully. However, they do not believe that *specific* familiar contexts have inherent moral relevance; insofar as people are free to choose, adjust and reconsider their familiar contexts, there is no need in principle to grant protection to specific contexts just because, for instance, they risk disappearing.²⁵

Underlying Kymlicka's conviction is a normative principle that classical liberals would reject, namely, that of the equal prosperity of (minimally reasonable) 'societal cultures'. The principle holds that a society should be so organised so to ensure that 'societal cultures' have an equal chance to flourish and do not risk perishing due to their diminished capacity to attract new adherent or maintain the old

²⁴ The idea that multiculturalism is an unnecessary addendum to liberal theory, for the latter already guarantees everyone's freedom to choose a cultural context, has been already defended by, among others, Ronald Dworkin, Jeremy Waldron, and Chandran Kukathas (Dworkin, 1989), (Waldron, 1992) (Kukathas, 1992).

²⁵ In the recent debate, Chandran Kukathas has probably provided at the same time the most comprehensive and the most radical defence of this position, which he summarises by writing that the principles of a free society 'should take as given only the existence of individuals and their propensity to associate; they need not and should not assume the salience of any particular individuals or of any particular historical associations (Kukathas, 2003, p. 4).

ones. Many state decisions will have an impact on the prosperity of 'societal cultures' and complete neutrality is in many regards infeasible as the state cannot refrain from adopting such things as a national language or legislation that incorporates some cultural values. Hence, the state must ensure, through such measures as the attribution of cultural rights and direct interventions in defence of vulnerable 'societal cultures', that the principle of equal prosperity is respected.

All 'societal cultures' equally deserve this form of recognition on the part of the public authority because their disappearance would not merely be an accident in someone's wellbeing. The disappearance of a 'societal culture' would prevent its previous adherents from continuing to exercise their freedom, as 'freedom involves making choices amongst various options, and our societal culture not only provides these options, but also makes them meaningful' (Kymlicka, 1995, p. 83). In a similar vein to Kymlicka, Alan Patten, another advocate of cultural rights, argues, about members of a 'disappearing culture', that '[t]heir options are significantly worse when their culture fares poorly than they are when the culture survives and flourishes' (Patten, 2014, p. 70).²⁶

The solution for both multicultural liberals (Galeotti, Kymlicka, Patten, Taylor), or anti-liberal identity theorists (Young), is that the political institutions engage in a policy of 'hands-on', direct recognition of the differences existing in society, which is the exact opposite of 'hands-off', traditionally liberal, neutrality.²⁷

²⁶ Despite agreeing with Kymlicka on this point, Patten ultimately rejects the idea that the preservation of cultural contexts can provide a good normative ground from which to justify the existence of cultural rights.

²⁷ I partly owe this distinction to (Balint, 2015).

4.2.3 The Problem of Evaluations and the Failure of 'Hands-on Recognition'

The question is whether 'hands-on recognition' respect can ever be taken as a plausible expression of the ideal of equal respect; my understanding is that it cannot. The reason behind my conviction is that 'hands-on recognition' gravely and irremediably compromises the universality constraint that, I argued before, is an inescapable component of any plausible account of equal respect.

Advocates of 'hands-on' respect acknowledge its symbolic role in providing an official recognition, on the part of the state, that a particular doctrine is legitimately part of the fabric of society. Could this recognition be independent of evaluations? It seems that the answer can only be negative; in order for the recognition to come about, the authority needs to determine whether the doctrine under analysis is, at the very least, not publicly detrimental. The evaluation would serve the crucial purpose of avoiding the public recognition of repugnant views, but it might also prevent the conceptual mistake of attributing respect to the wrong kinds of entities, such as, in Kymlicka's model, conceptions of the good that are not comprehensive enough to qualify as 'societal cultures'.

As 'hands-on' respect depends on evaluations, one may find that I already have a reason, derived from the previous observations, to reject this model. I do not take this point to be very compelling, as the Kantian ideal does not in fact bar *all* possible kinds of evaluations about the objects of equal respect. As I wrote before, the evaluations that the hypothetical attribution of 'as-if rationality' excludes are those that concern people's use and possession of their capacity to form and revise a conception of the good or a sense of justice. However, other evaluations are unavoidable. We might need evaluations, for instance, to determine the age at which average human beings acquire a capacity to form and revise a conception of the

good; other might be introduced (although I remain agnostic on this point) to adjudicate on whether some non-human animals are somehow capable of conducting choices following presumed values.

The evaluations that Kymlicka's account presupposes might be seen as analogous to those used to determine whether some human being has reached the age or cognitive development needed for the attribution of 'as-if rationality'. But the analogy is misleading. In Kymlicka's case, entire conceptions of the good, and not individuals, are subject to evaluations.

For example, we might say that political ideologies, under Kymlicka's model, are not entitled to any form of respect that is due to 'societal cultures', regardless of their content, because they are not fully comprehensive. This stark separation between 'societal cultures' and less than fully comprehensive packages of ideas seems dubious; it is far from obvious that one can always distinguish conceptions of the good that provide their holders with the values necessary to make meaningful choices from ones that play a more trivial role in people's lives. Some political 'ideologies', as opposed to mere political affiliations, tend indeed to overcome the traditional boundaries of political commitments and influence their followers' lives to a much wider extent.

Kymlicka's reply in defence of his isolation of 'societal cultures' as the only ones deserving recognition could be that, despite being more comprehensive than average political affiliations, political ideologies are still not central enough to someone's identity. The idea would be that a political ideology, however comprehensive, is an affiliation that people adhere to in virtue of the fact that they already possess in its integrity a capacity to make significant choices in life. People who are forcefully deprived of their 'societal culture' would by contrast lose irremediably their ability

to exercise freedom of choice, as they would have been deprived of the values and criteria that they use to take all meaningful decisions.

This point becomes controversial, however, when one starts considering some of Kymlicka's favourite examples. Would French-speaking Canadians, for instance, lose their capacity to exercise freedom meaningfully in case the preservation of the French language were not so safely guarded any longer in Québec? That seems doubtful. Let's take 'meaningful' choice to be those that are *relevant* for the person who takes them (the person would never accept to see the matter settled through a coin toss) and that do not concern matters of mere biological survival. 'Trivial' choices are those that, by contrast, are *irrelevant* for the subject or concern matters of mere biological survival. Now, even in the worst-case scenario, where the entire French culture in Canada would disappear, formerly French-speaking Canadians would never be reduced to individuals incapable of taking meaningful choices.

By contrast, people who are forced by the public authority to abandon their political affiliation might understandably feel that, with the deprivation of their political representation, they have entirely lost their ability to take decisions in the political arena that might promote their interests and values. They might feel, in other words, politically estranged from their own society and therefore incapable of having a say in any political decisions that might affect them. In their case indeed elections might feel like coin tosses or, in case one party is marginally less prejudiced against their instances than any other, a mere matter of 'political survival'.

These points against Kymlicka are all, I believe, significant, but they are based on a concession to him. Specifically, I have uncritically conceded that an official recognition, on the part of the state, of specific cultural contexts, is morally unproblematic. I now want to contest that concession.

To provide an example of the kinds of state recognition that characterises Kymlicka's model, let's suppose that the state decides to assign a special form of protection to specific 'societal cultures', following a reasonable prediction that, without that protection, those will eventually disappear. In order for this policy to succeed, however, the state has first to decide which 'societal cultures' deserve to be preserved, and which ones, on the contrary, can be legitimately left to die out. For instance, the state could decide that the only conceptions of the good that deserve an official recognition and that are entitled to protection in case a credible threat for their survival comes about are those that meet certain criteria of public acceptability. Excluded conceptions might be those that exhibit discriminatory attitudes outside or inside the group membership.

Notice that, if the state did not engage in these kinds of adjudications, a policy of 'hands-on' protection for conceptions of the good would be infeasible. I therefore take issue here with the view of those who believe that 'symbolic recognition [...] is not a scarce commodity, posing problems of distribution'.²⁸ When 'symbolic recognition' is meant to entail that the recognised view deserves a form of protection designed to guarantee its prosperity, then indeed it turns into a scarce resource, and distribution problems arise. Not only the state has not enough resources to guarantee that all 'societal cultures' equally prosper, but a further complication is that requests from protection often take place against the public expression of other, supposedly oppressive, 'societal cultures'. Take the insidious problem of the public expression of discriminatory views, and consider in particular the case of discriminatory views that are parts of a broader conception of the good such as, for instance, a political ideology. Now, it appears that in this case at least the public authority dis-

²⁸ (Galeotti, 2002, p. 16).

tributing 'symbolic recognition' has to make a choice. It can either safeguard absolute freedom of expression, and thus frustrate the demands of protection coming from discriminated groups, or take some measures against the public expression of the discriminatory conception of the good. In the end, the protection accorded by the state to some conceptions of the good will serve to defend the latter against the proliferation of other views that the state itself, rather than protecting, takes active measures to contain.

We might then ask; what does the state communicate when it decides to accord recognition to a certain 'societal culture' and not to others? My answer is premised on the idea that conceptions of the good play a significant role in both the individuation, on the part of an external observer, and the self-definition, on the part of members, of 'societal cultures'. To exemplify, an external observer will identify Muslim believers as those that hold a common religious conception, and similarly a Muslim person will self-identify as such as long as she maintains her commitment to certain elements of the Muslim faith. What is trivially true of a purely religious community can also be shown to be correct in cases of communities where religion does not play such an irreplaceable identifying role. Jewish communities, for example, have traditionally included both religious believers and agnostics, but even non-believing Jews must hold firm a particular belief in order to maintain their affiliation to the Jewish community, namely the belief that they see themselves as (secular) Jews.²⁹ Alongside that single belief about her own self-identification, the person committed to the view that she is still a member of a community must also find her-

²⁹ The same applies to self-identified 'secular Muslims'; the example with Judaism was just made more obvious by the usual characterisation of Judaism as constituting a people, other than a religious community.

self in agreement with at least the most significant values that the community embodies. Were it not this case, she would have no interest in remaining part of the community; let alone, she could be accused of apostasy and forced to leave.

Adherence to a specific culture is in fact one of the most primitive and frequent ways for every person to exercise the capacity to form and revise a conception of the good; virtually every person sooner or later in life is subject to the choice of some form of affiliation. It follows that states according recognition to a particular 'societal culture' and not to others would communicate that only the members of the first sociocultural group have used their capacity to form and revise a conception of the good in a politically acceptable way.

In the recognition respect model, therefore, only the holders of the view that somehow *deserve* to be officially recognised by the state, in virtue of their being at least publicly acceptable, are entitled to recognition respect. The Kantian contention that we ought to 'respect a human being even in the logical use of his reason' is replaced by the appeal to respect human beings only in virtue of the extent to which their views contributes at least non-negatively to the public debate.

Why, one might ask, is this a problem in the first place? The rationale behind 'hands-on recognition' is that of safeguarding minority groups that suffer from marginalisation in society, despite their holding views that pass the threshold of public acceptability (they do not discriminate inside the group and do not hold discriminatory views against people outside the group). This is a noble intent, and my critiques might appear excessively harsh.

But the original intent of an ideal of equal respect, I argued previously, was that of identifying the treatment that, in a particular context, all persons are entitled

to, purely in virtue of their personhood. This is the view I defended before by observing how universality is a component that seems to be germane to the very existence and significance of the notion of equal respect.

Any account of respect that leads to the rejection of universality renounces the project of equal respect altogether. It accepts as an unfortunate consequence of pluralism that, because certain views will inevitably collide with whatever standard of public acceptability one decides to adopt, then those who endorse such views are not entitled to fully equal respect. Specifically, the holders of such publicly unacceptable views can still receive respect in other sectors of the public agenda (for example, when it comes to the distribution of resources that are collectively owned or the assignment of basic rights), but not in their capacity to form and revise a conception of the good.

To further strengthen this point, it is worth drawing a comparison with one influential way of defending freedom of expression. The account I am considering has been proposed by, among others, Joseph Raz and centres on the idea that the 'public portrayal and expression of forms of life validate the styles of life portrayed' and that 'censoring expression normally expresses authoritative condemnation not merely of the views or opinions censored but of the whole style of life of which they are a part' (Raz, 1995 b, p. 153).³⁰ Although I have not used the language of validation, the points I have stressed have a lot in common with Raz's account of freedom of expression. In the same way as, according to Raz, being able to publicly express

³⁰ Raz's validation-based account is, even for Raz himself, just one possible way to defend freedom of expression. It is however interesting to notice how my critique of the recognition approach bears some resemblance even to another influential defence of freedom of expression, Thomas Scanlon's autonomy-based account. Scanlon argues that 'the harm of coming to have false beliefs is not one that the autonomous man could allow the state to protect him against through restrictions on expressions' (Scanlon, 2003 b, p. 17). Similarly, I would say that the harm of being exposed to a possibly disrespectful culture is not one that autonomous people could allow the state to protect them against through restrictions on the types of values and cultures that the state can officially recognise.

something one considers of value amounts to a public validation of one's lifestyle, so, under my account, being free to pursue the values one has chosen amounts to a formal recognition that one is using his capacity to form and revise his conception of the good in a publicly acceptable way. *Vice versa*, in the same vein as, for Raz, public acts of censoring 'not only deny the individuals the opportunity to provide validation but constitute official condemnation of the ways of life aspects of which are censored' (p. 158), so, in my view, the official denial of recognition respect to a particular culture amounts to a public condemnation of the way in which members of that culture have exercised their capacity to form and revise a conception of the good.

Furthermore, 'recognition respect' is an account that only solves, practically speaking, a part of the problem. Indeed, although it has a lot to say on the appropriate treatment of discriminated and stigmatised communities, it is silent on the equally significant question of the right treatment that is due to those members of society that do not hold publicly acceptable views. Granted that the latter views do not deserve any formal protection, can they at least be tolerated or should the state actively take measure to 'contain' them? And are the holders of unreasonable views somehow culpable and therefore punishable for that? All these are questions that 'hands-on' recognition accounts leave unanswered.

4.3 Scepticism, 'Mediated Respect' and 'As-If' Rationality

The difficulty in according equal respect to citizens in circumstances of endemic and irreducible disagreement has led some authors to express overall scepticism about the use of the language of respect applied to conceptions of the good.³¹ Among these sceptics, Peter Jones has developed an account of 'mediated respect' (or 'mediated

³¹ See especially (Jones, 2006) (Jones, 2015) and (Balint, 2006) (Balint, 2015).

recognition') for differences, which, in the words of a commentator, 'involves initially respecting something as a general category, for example as a person or citizen, and then as a consequence of this respect, respecting the specific difference' (Balint, 2006, p. 53). Conceptions of the good life are accorded respect purely as a consequence of the subjective importance they have for the agents who are respected as citizens. The main advantage of this proposal, according to its advocates, is that 'mediated respect' for differences would be consistent with a tolerant perspective in which the doctrines object of respect are devoid of any positive evaluation.

This last aspect fits easily with the independence from evaluation I isolated before as a component of the Kantian ideal. Moreover, 'mediated respect' seems also capable of providing plausible answers to a vaster array of questions than 'hands-on recognition respect'. Its use, for instance, is not restricted to the case of marginalised communities holding publicly acceptable views.

A significant drawback of the proposal, however, is that its own advocates are sceptical about the extent of its successful implementation. Peter Balint, for example, seems doubtful that this approach could be used between citizens, for 'the radically multiculturalised citizen', the one capable of according respect to ideas he or she ultimately despises, 'is surely an exception, even though he or she might be the ideal citizen for many' (2006, p. 55). Peter Jones, who introduced the label 'mediated recognition' in the first place, shows to be even more sceptical about the possible use of the language of recognition respect if applied to the context of doctrinal pluralism. He believes that the main value a liberal democratic order ought to promote in circumstances of endemic pluralism is that of toleration, which is realised through 'an order that secures freedom of religion for its citizens by not allowing

them to use political power either to privilege their own faith or to suppress the faith of others' (Jones, 2015, p. 546).³²

In a similar vein, Chandran Kukathas suggests that, because 'not all groups and practices are in fact respectable' (Kukathas, 2003, p. 33) – presumably in the 'appraisal' sense of respect – then an 'easy tolerance of differences into whose character we choose not to inquire' (p. 34) is what both morality and prudence require.³³ In those conditions where the projection of competing views can easily escalate into a conflict, Kukathas argues, a 'live and let live' approach is preferable to the deep investigations into a society's various cultural values that the recognition approach mandates. The point seems to align perfectly with my argument so far; multicultural societies are much more diverse and complex than multiculturalist theorists usually interpret them and assigning special forms of recognition to certain cultures is not a politically neutral move. According to Kukathas, therefore, one should focus instead on what can be done to maintain the political order and guarantee everyone the respect of the fundamental human right to associate with whomever one wants and its two corollaries, the 'freedom of disassociation' and the 'mutual toleration of associations' (p. 4).

In describing these kinds of theories as sceptical, I am not accusing them of retreating to a kind of Hobbesian pessimism regarding the possibility that fundamental rights might be given protection when conflicts are in place. The scepticism, by contrast, concerns the idea that participants to a value-laden public discussion might receive anything more than a generic entitlement to express their view without impediments.

³² See also (Jones, 1994) and (Jones, 2006) for similar reflections about toleration.

³³ I will have more to say on Kukathas's model of toleration in the next chapter.

My aim is now to show, starting again from the Kantian ideal, that such scepticism is misplaced, and that in cases of disagreement we are not morally required to retreat to a policy of mere toleration, when this is understood as the restraint from directly taking measure to contain the expression of views one finds despicable.

The point of departure is, now unsurprisingly, the Kantian ideal. An interesting suggestion is offered by Kant when he recommends, speaking of the agent one is disagreeing with, 'to suppose that his judgement must yet contain some truth and to seek this out'. The expression 'contain some truth' is clearly idiomatic but seems to imply that, when one intends to engage with disagreement in a respectful way, the judgement one disagrees with cannot be shrugged off as a mere absurdity.

As I argued before, one way of adopting the Kantian attitude of 'supposing that his judgement must yet contain some truth' is to attribute to others some form of rationality. Conceiving others as capable of forming a conception of the good and a sense of justice means interpreting the views they advance as the product, however twisted and mistaken, of a faculty we all equally possess. The commonality of the source further implies that those views are never, at least in principle, something that we cannot potentially engage with, although sometimes this will require a considerable effort. Treating others who disagree with us with equal respect demands taking their views as at least potentially *intelligible*, in a sense that I will now explain.

To begin with, it is worth striking a comparison with a typical attitude in cases of endemic disagreements, particularly those touching matters that are of utmost concern to the participants, that of individuals merely *talking past* each other. Underlying this attitude is probably the belief that nothing the other party has said or could ever say is worthy of engagement as it is just the product of prejudice and stubborn attachment to unquestionable idiosyncrasies. The positions displayed by

the other party are not even granted the status of *arguments* and this alone prevents any further development of the discussion. On the contrary, if one starts seeing that, beyond the prejudices and idiosyncrasies (which might be present even in one's side) the other party is advancing *arguments*, then that move alone can give a positive turn to the discussion; unlike prejudices, arguments can be countered, their assumptions exposed, their empirical components reconsidered, their logical conclusions showed to be flawed. The discussion has then been turned from a mere clash of incompatible contentions to a debate that has at least the potentiality of leading to sharable conclusions.

Now imagine we take a step further and we assume that not only the other parties are advancing arguments that can be engaged with, but that the values that they seemed attached to, and that initially appeared the most formidable obstacle to the development of the discussion, could now be seen as products of a faculty we all share. That means that even values, alongside argument, can be engaged with, because they too are, at least potentially, intelligible. What we can understand, in particular, is the *relevance* of that value for the person who is attached to it, even when we would never personally be moved by it. We understand the relevance because we see that these values, however bizarre or even repugnant may appear to our own current evaluative system, are the result of an urge we all feel, which is that of interpreting the world surrounding us in a way that is not merely instrumental to our biological conservation.

What does intelligibility imply, however, when applied to values? In the final chapter, I will show the implausibility of one possible interpretation, which is that we ought to let the subjective relevance of others' commitment be a direct constraint on our moral conduct. Here, however, I want to suggest a more minimalist normative proposal, which is that, in circumstances of public deliberation, we show equal

respect to others only if we prove to be willing to make sense of their position in the most unbiased possibly way.

I now intend to substantiate equal respect in circumstances of public deliberation through the formulation of negative and positive moral duties.³⁴ The positive duties derived from the ideal of equal respect would correspond more neatly to the Kantian requirement of supposing the other party's judgement 'must yet contain some truth'; agents would discharge them by trying to attain an unbiased comprehension of the position they contest. Symmetrically, the negative duties correspond to the Kantian suggestion of refraining from 'calling them absurdities, poor judgement and so forth'. Notice that 'not calling them absurdities' need not be taken literally; the negative duties surrounding equal respect do not prevent us from producing harsh value-judgement concerning others' publicly advanced views. What they do imply, however, is that a position advanced in a debate conducive to the enactment of a public norm deserves a basic, minimal form of respect, which is that it

³⁴ The definition of a moral duty falls beyond the borders of the present work. Suffice it to say, however, that I take as deeply relevant the distinction, sometimes neglected, between a moral duty and a (moral or legal) obligation. Two features are particularly striking in drawing a comparison: an obligation differs from a duty both for its specific content (it demands the performance or omission of a specific course of action) and for its universal validity; an obligation is binding, when it is, whatever the circumstances. This is shown by the fact that not complying with an obligation immediately generates a further obligation, which is that of providing a justification (or an excuse). The distinction between duties and obligations has a long philosophical pedigree. H.L.A. Hart distinguishes them in (Hart, 1955, p. 179) by claiming that '(1) obligations may be voluntarily incurred or created, (2) [...] they are *owed* to special persons (who have rights), (3), [...] they do not arise out of the character of the actions which are obligatory but out of the relationship of the parties' (footnote 7 – emphasis in the text). Thus, for Hart I might have a (moral) *duty* to respect promises (in general), but a (special) *obligation* to water my neighbour's plants once a day, as I promised to him. That the notion of duty always entails some sort of *prima facie* (or, in the current vocabulary, *pro tanto*) character is further recognised by R.B. Brandt in (Brandt, 1964, p. 377). Finally, Rawls does not distinguish duties and obligations in the most straightforward way, but, in TOJ §51-52, he clearly follows Hart in treating separately (a) a 'natural duty' 'to comply with and to do our share in just institutions when they exist and apply to us and [...] assist in the establishment of just arrangements, when they do not exist' (pp. 293-294) and (b) an 'obligation to do [one's] part as specified by the rules of an institution whenever [one] has voluntarily accepted the benefit of the scheme or has taken advantage of the opportunities it offers' (p. 301).

must be taken for what it is to the eyes of its holders, and not deceitfully misinterpreted.

The manipulation of another's view, even when conducted following a noble aim, such as an attempt to identify a compromise, would violate the moral requirement to treat every person as an end in itself; specifically, it would betray an attitude of not considering the other as a person capable of forming and revising an intelligible conception of the good. The only way to ensure that the deceitful reconstruction has not occurred is by showing that the agents are willing to make sense of each other's position, in the most unbiased way.

Two specifications are now in order. Firstly, the label 'making sense of each other's position', which I have previously used, does not necessarily lead to condoning or excusing views one disagrees with; although we might expect that mutual comprehension will often result in tolerance (*tout comprendre c'est tout pardonner*), the opposite upshot, where one or the other party reinforces the belief in the perniciousness of the other's view, is equally possible. Secondly, the attitude must not be confused with the outcome of the process; nothing in the account I have so far defended implies that only the state of full mutual comprehension would mark the expression of mutual respect.

Going back to a previous example about the abortion controversy³⁵ will help clarify what I mean by a willingness to understand others' positions in the most unbiased way. Some agents in the example were suggesting compromise positions about the abortion debate that I have shown were ultimately unattractive to the supporters of the 'pro-life' party. The advocates of these 'sham compromise' positions do not have to be motivated by a malicious attitude; they might be genuine in their

³⁵ See p. 161.

desire to reach an agreement on such a disputed question. But their proposals reflect a (more or less conscious) refusal to understand the full significance of the 'pro-life' position; hence, the 'pro-life' supporters are in fact treated as a mere means to the (noble) end of striking a compromise, and not as persons capable of holding and living up to an intelligible conception of the good.

This conclusion might look unnecessarily fatalistic; acknowledging that the 'sham-compromise' positions are unacceptable to the eyes of the 'pro-life' might make the attainment of an agreement among the competing parties look even less feasible. Someone could plausibly see my suggestion as leading to paradoxical results; on the one hand, I recommend equal respect as a normative constraint on the achievement of agreements in circumstances of public deliberation, and, on the other, I make the realisation of those agreements less feasible. But this allegation is based on an illusion, namely that, by slightly adjusting one's own position, without even considering the position the other party is really advancing, compromise is made easier. As argued above, it is unlikely that an agreement could be reached through a compromise position that in fact considerably leans towards favouring one's position, and that, most importantly, ignores the adversary's viewpoint.

Even if the possible objector were right, however, in thinking that the Kantian ideal of equal respect would make the achievement of a compromise less feasible, I still believe we have good reasons (at least good *pro tanto* reasons) to promote equal respect. I believe, in particular, that if a compromise were reached through the misconstruction or manipulation of another's view, even when this is done with a benign attitude, something would be missing on a deontological perspective.³⁶ The result, however welcome for civil harmony, would have come about at the expense of

³⁶ I take for granted that a generally deontological perspective underlies the deployment itself of the normative category of equal respect.

some citizens' possibility to exercise their capacity to form and revise a conception of the good.

The minimalist character of my proposal should be evident; I now conclude by signalling out its advantages. My account is an interpretation of the ideal of equal respect that I believe is at the same time consistent with the features I identified as definitive of the ideal and practically informative. I call it 'practically informative' because I see it as capable of informing practical norms, albeit of a minimalist kind, namely, norms that mandate that each party demonstrate its commitment towards the most unbiased understanding of others' positions. The norm is of moral significance because it represents the only way to show respect towards others' capacity to adopt and live up to a conception of the good. It is also, however, of prudential significance, because only an unbiased understanding of the others' positions can prevent the advancement of 'sham-compromise' proposals that can in fact trigger an escalation of the debate.

If my account is viable, it is therefore possible to find a role for the Kantian ideal of equal respect in circumstances of endemic value-disagreement, and, most importantly, the retreat to 'restraint-tolerance' is made unnecessary.

5. Stability as Toleration

The previous chapters have presented a series of answers to the question I started with, namely, why is value-disagreement problematic in political and moral theory. I have tried to show that, the more we take disagreement seriously, the harder it gets to arrive at a plausible account of legitimacy. By contrast, my conclusion from analysing the question of equal respect were overall positive; there is a non-trivial interpretation of the value of equal respect that we can use to guide our treatment of our fellow citizens in the public forum and that may successfully take into account their differing from us in the end they intend to pursue.

There is no doubt, however, that the notion that has been most central in the debate on disagreement is that of political stability, not else because of the prominent role it plays in Rawls's Copernican revolution in the years following the publication of *A Theory of Justice*.¹ This centrality plausibly derives from the inevitability of seeing disagreement as a constraint on stability. Even the Platonist who believes that the reality of ideas is wholly independent of people's beliefs, and for whom therefore disagreement between people is a philosophically negligible state of affairs, has to recognise that, if justice matters, then the stability of a just state of affairs matters too, and, insofar as people disagree about what justice is, the stability of justice is going to be a costly enterprise.

We can initially define the problem of stability as the problem of determining which costs are worth undertaking when trying to maintain a particular just state.

¹ The Copernican Revolution I am referring to is described by Rawls himself in the Introduction of *Political Liberalism*, when he writes that '[t]he fact of a plurality of reasonable but incompatible comprehensive doctrines – the fact of reasonable pluralism – shows that, as used in *Theory*, the idea of a well-ordered society of justice as fairness is unrealistic'. And there he significantly adds that '[t]he account of stability of a well-ordered society in part III [of *Theory*] is therefore also unrealistic and must be recast' (PL, xvii).

One solution to the problem of stability emerged in the Renaissance with the works of Machiavelli and Hobbes and gained the name, which will come back later in this chapter, of 'political realism'. Its slogan is *extra rempublicam nulla justitia*, meaning that, being the institution of the state the principal prerequisite of justice, *any* moral cost is worth undertaking when the preservation of the state is at stake, and the question will merely concern the best strategy that the *Principe* or the Leviathan needs to follow in that exercise of self-preservation.

I tend to believe, following Samuel Freeman's reconstruction (Freeman, 2003), that this is not the question Rawls was interested in when he started conceptualising political stability in *Theory*. For once, if there are moral costs inherent to the preservation of justice, they cannot compromise the fundamental constraint that the state must be legitimate, because, as Locke had already recognised, were this condition relaxed, then the state would not even be preferable, morally and prudentially, to its negation. Freeman suggests Rawls is interested instead, in the presentation of stability in *Theory*, in a problem that particularly troubled Kant,² namely the one about the reconciliation between, or in Rawls's words the 'congruence' of, the social right and the individual good.

But the demonstration of the 'congruence' in *Theory*, which I will reconstruct in the last section of the present chapter, depends on a substantial disregard for the existence of pluralism. What Rawls merely intends to show is that, within a plausible conception of rationality, the full development of a sense of justice is an intrinsic, almost natural, good, so that, to quote Freeman, 'the sense of justice and the desire to express one's nature are "practically speaking the same desire"' (p. 296).³

² My reference is to the 'antinomy of practical reason' Kant mentions in the *Critique of Practical Reason* (Kant, [1788] 1999, p. 231), the problem of the reconciliation between morality and happiness to achieve the highest good.

³ Freeman is quoting from TOJ, p. 501.

Rawls's Copernican revolution in the years following the publication of *Theory* rejects that tacitly perfectionist component of his theory.⁴ The problem then becomes how to demonstrate that a reasonable theory of justice can be compatible with a *plurality* of equally reasonable conceptions of the good. Once that has been shown to be possible, the problem of stability is solved by providing an account of 'stability for the right reason'. The account is based on the idea, presented in detail in the chapter on public justification, that various reasonable conceptions of the good can reach a consensus on basic principles of justice that a state can appeal to in justifying its authority.

I have already significantly relied on the Rawlsian notion of 'moral personality' in the previous chapter. My point was that attributing to our fellow citizens a capacity to organise an intelligible conception of the good is both a pre-condition and an expression of our treating them with the particular type of respect that is appropriate to the circumstances of public deliberation. In the present chapter, I am going to similarly defend, so to speak, the Rawls of *Theory* against his most mature self in *Political Liberalism*. Although the argument will not be grounded in Rawls himself, I will argue that the problem of stability only finds a solution, odd as it might appear to a neutralist liberal, in a kind of perfectionist ideal. More precisely, I will devote my attention, in a vein altogether similar to the previous chapter, to the type of *attitude* that citizens need to display in order for political stability to be realised in circumstances of justice.

The attitude I will defend is one of 'toleration'. As I will specify later, I do not defend my proposal as the most correct account of toleration, but only as a plausible use of the term. The attitude depends on a recognition of fellow citizens in society as 'just as entitled as we are to contribute to the definition of our society', a definition

⁴More on Rawls's minimal perfectionism in section 3 in the present chapter.

I have derived from a prominent account of toleration (Scanlon, 2003 c, p. 193). The attribution to others of this moral status, and the acknowledgement that they are as entitled as we are to shape society according to some of their values (but not, as I will show, to shape society *as they see fit*), provides precisely that public assurance that neither 'public reason' nor 'convergence' nor 'political realism' can provide. For, if I accept that others might condition the way my society looks like, I will not refuse to comply every time my own attempt at shaping society has not been realised.

Absent that disposition, I have simply no reason, moral or prudential, to keep complying with the common rules whenever they require that I partly compromise the pursuit of my individual end. Which means that I will only comply when all-things-considered moral reasons require me to do so or when acting that way helps promote my self-interest. Absent that disposition, in other words, stability only obtains under two conditions: perfect justice or perfect coercion.

Before I defend my account in the final section of the present chapter, I will prove why the attempt to defend political liberal public justification as a good model of stability fails. It fails, I will argue, because it either demands citizens too much, or it provides no assurance at all. As the argument I will present there mainly builds on my conclusions in the chapter on public justification, that section will be significantly brief.

A more interesting challenge, on the other hand, is advanced by contemporary 'political realism'. Against Rawls and other alleged 'political moralists', contemporary 'political realists' try to produce a political theory that is not merely an application of morality to the realm of politics. I will argue that the main insight we can derive from political realism concerns the limits and ambitions of normative political theory; insofar as political theory comprises elements of idealisation and abstraction, it cannot aspire to produce accounts that are immediately applicable as

action-guiding norms. However, this type of objection, which is particularly insidious against the proposals of scholars who aim to use the conceptual instruments of political theory to derive a universally applicable account of political stability, has no bite against my proposal. As my intent is that of pointing out a type of attitude that, when widespread, might enhance the stability of a just constitutional order, my ambitions in terms of the application of the ideal to concrete reality are already limited. Which does not mean, however, as I will specify in the closing remarks, that my account is devoid of any possible practical implications, but that it does not assume that action-guiding normative conclusions can be immediately derived from it. How indeed a spirit of toleration can be made to prevail in society, so that the society itself can aspire to be stable, is a matter that, as the realist insight rightly reminds us, will always depend on contextual minutiae that a theoretical account will be unfit to capture in their integrity.

5.1 Public Justification and Stability – A Sceptical Perspective

5.1.1 Stability as a Problem of Assurance

Whilst defending Rawls's public reason as providing a good account of political stability, Paul Weithman distinguishes Rawls's peculiar take on the question of stability from the one that has been prevalent in the history of political thought. He writes that, whereas most political thinkers have been concerned about how to *impose* stability, Rawls was interested in how to show that a conception of justice is *inherently* stable: 'a conception of justice is inherently stable if a society that is well-ordered by it generally maintains itself in a just general equilibrium and is capable of righting itself when that equilibrium is disturbed' (Weithman, 2011, p. 45).

To illustrate the difference, we can imagine two different solutions to the classic free-rider problem. Take a scheme of cooperation that distributes benefits to

contributors and non-contributors alike, and where the possible sanctions used to deter people from non-contributing either do not exist or are particularly ineffective. Then, the obvious self-interested rational response for all agents involved is that of not contributing at all, for at least two reasons. The first reason is short-sighted and morally perverse, but (unfortunately) quite ingrained in human nature. It tells agents that they should not contribute, whenever they can receive the benefit anyway – let others be suckers! The second reason, on the other hand, is not blatantly immoral and is what ultimately demonstrates the rationality of non-contributing. Assume that each actor has realised that, were everyone else to non-contribute, the entire scheme would go bankrupt and there would be no benefit. Even though this conclusion provides a good reason to contribute, each actor has no assurance, in the present circumstances, that others will in fact contribute, thus maintaining the scheme operative. Hence, contributing does not guarantee that the scheme will not go bankrupt anyway. Even considering the long run, therefore, self-interested rationality suggests to each agent the socially detrimental conduct of not contributing.

We can illustrate this through a simplified two-player game, which we can call ‘the free-rider dilemma’. The parenthetical quantification refers, as usual in game theory, to positive or negative variations in each agent’s utility.

THE FRE-RIDER DILEMMA	Player B contributes	Player B does not contribute
Player A contributes	A and B: full benefit received (2) + contribution (- 1.5) = 0.5	A: partial benefit (1) + contribution (- 1.5) = -0.5 B: partial benefit = 1

Player A does not contribute	A: partial benefit = 1 B: partial benefit (1) + contribution (- 1.5) = -0.5	A and B: no benefit received = 0
-------------------------------------	--	----------------------------------

Now, the immediate solution to the free-rider dilemma is to introduce those sanctions that were characteristically absent or inefficient in the previous scenario, so that, once the cost of sanctions are factored in, the new game becomes:

LEVIATHAN	Player B contributes	Player B does not contribute
Player A contributes	A and B: full benefit received (2) + contribution (-1.5) = 0.5	A: partial benefit (1) + contribution (- 1.5) = -0.5 B: partial benefit (1) + sanction (-2) = -1
Player A does not contribute	A: partial benefit (1) + sanction (-2) = -1 B: partial benefit (1) + contribution (- 1.5) = -0.5	A and B: no benefit received (0) + sanction (-2) = -2

The problem is that the Leviathan solution is characteristically illiberal. Notice that, in order to move the Nash-equilibrium from the bottom right to the top left square, the sanctions introduced to deter agents from not contributing have to exceed, in the fictional quantification I have used, the absolute value of the benefit received. Deterrence of sanctions is usually interpreted, starting with the pioneering

Enlightenment contributions of Bentham and Beccaria,⁵ as a function of two variables, namely the likelihood of being caught in the act of legal evasion (the ‘certainty’ of the penalty) and the severity of the sentence itself. Regardless of whether the Leviathan, in its attempt at producing effective deterrence, prioritises the severity aspect, or the ‘certainty’ one, or gives equal consideration to both, the penal system resulting from that choice is going to be characteristically illiberal. It will be a penal system composed of disproportionately severe sanctions, or of extreme interferences into people’s private lives, or a combination of the two.

The liberal solution would consist in bringing about the same result, namely a condition whereby the rational conduct for each participant is to contribute, without relying (at least not exclusively) on a coercive mechanism such as punishment.⁶ The liberal game would be one in which the agents believe that contributing to the scheme is not only the right thing to do in an abstract moral sense but also the thing that they feel most inclined to do, because it does cohere, at least on an ordinary basis, with the pursuit of their individual good.

This ‘liberal benefit’, which I am going to quantify in the next game, is what Rawls would think possible to obtain, in *Theory*, through the realisation that the cultivation of a sense of justice constitutes an intrinsic good and, in *Political Liberalism*, through a public justification that relies on an ‘overlapping consensus’ of reasonable conceptions of the good. The process whereby the liberal benefit comes about is different, but the result is similar. In the first case, the ‘liberal benefit’ is a good in itself, that adds to the more mundane general benefit (presumably of an economic form) that derives from the use of a common-pool resource. In the second case, the ‘liberal benefit’ is not a good in itself, but is an offhand expression used to describe the

⁵ (Bentham, 1970 [1780]); (Beccaria, 1995 [1764]).

⁶ Sanctions are of course just an example of a coercive mechanism. We can well imagine a Leviathan that does not need sanctions at all because it can rely on extremely efficient coercive tools.

recognition, on the part of each agent, that the scheme of cooperation she is required to contribute is one that, in the terms of her own reasonable conception of the good, she finds acceptable. Regardless of its origin, the emergence of the liberal benefit will imply that ‘when agents’ preferences are transformed, such as that they internalise moral [...] standards and thereby come to desire to act fairly in a regime of reciprocal fairness, the costs and benefits of compliance and enforcement are similarly transformed’ (Hadfield & Macedo, 2012, p. 44). Weithman puts it in more explicitly game-theoretic terms, by affirming the existence of a *Political Liberalism’s* ‘Nash claim’, holding that, once the overlapping consensus has obtained, ‘each member of the well-ordered society judges, from within her comprehensive view, that the balance of her reasons tilts in favour of maintaining her desire to live up to the values and ideals of justice as fairness, at least when others live up to those values and ideals as well’ (Weithman, 2011, p. 302).

We can give a positive numerical value to this liberal benefit, so to derive a liberal game.

THE LIBERAL GAME (IMPERFECT)	Player B contributes	Player B does not contribute
Player A contributes	A and B: full benefit received (2) + contribution (-1.5) + liberal benefit (0.5) = 1	A: partial benefit (1) + contribution (- 1.5) + liberal benefit (0.5) = 0 B: partial benefit (1) + liberal sanction (-1) = 0
Player A does not contribute	A: partial benefit (1) + liberal sanction (-1) = 0	A and B: no benefit received (0) + liberal sanction (-1) =

	B: partial benefit (1) + contribution (- 1.5) + liberal benefit (0.5) =0	-1
--	--	----

The introduction of the liberal benefit has made the whole game indecisive. Every player knows that contributing is less risky, because only not contributing can generate the worst-case scenario where each of the players gets a negative benefit (-1). They also know, however, that, in case they do contribute and the other does not, the benefit they will incur will be numerically identical to the benefit incurred by the free-rider. Risk-aversion will therefore lead in the direction of unilateral contribution, but the natural desire not to be the 'sucker' in a scheme that distributes benefits to contributors and non-contributors alike will lead to the opposite conduct. The liberal game here introduced does not solve the 'free-rider dilemma'; whether it does depends on whether the actors involved are generally motivated from risk-aversion more than they are from the natural repulsion at being the sole contributors in a scheme that distributes benefits universally.

If we keep the value of the liberal benefit fixed, the only way the Imperfect Liberal Game can turn into a Perfect Liberal Game is to provide some assurance for the parties involved that their co-participants will indeed contribute. That way, we exclude the possibility that the other party is not contributing (or at least we render it statistically insignificant) and we automatically lead each party to prefer the contributing option as that will (1) lead to an overall increase in the benefit derived from the distribution of the common resource, (2) avoid the risk of incurring a sanction and (3) provide them with the liberal benefit. Now, the question worth posing is therefore whether the achievement of full public justification through either an

‘overlapping consensus’ or a ‘convergence’ mechanism can guarantee that kind of assurance.

5.1.2 Scepticism over Public Justification and Assurance

A number of Rawlsian theorists have tried to demonstrate that the use of ‘public reason’, that is, a set of ‘reasons [citizens] may reasonably give another when fundamental political questions are at stake’ (Rawls, 2005 [1997], p. 441) can provide each citizen with the assurance that their co-citizens will comply with public norms even when that might constrain the attainment of their individually chosen goals and purposes.⁷

The reasoning is simple, and *prima facie* compelling. By relying on an idea of public reason in their role as public officials or drafters of public norms,⁸ citizens show they can leave aside their idiosyncratic preferences in public, and thereby assure their fellows they will not stop cooperating the moment these preferences risk not being realised to the maximum extent. The assurance takes place through the public display, on the part of each participant, of a particular moral attitude, which,

⁷ See Weithman himself and (Hadfield & Macedo, 2012).

⁸ Rawls is adamant that public reason does not apply to the ‘background culture’, that is to say, the ‘culture of civil society’ (Rawls, 2005 [1997], p. 443) where conceptions of the good are formed and discussed. The limits in the use of public reason are however less restrictive than some in the literature (see Weithman himself or Kogelmann-Stich) tend to argue. For example, although public officials are the main actors involved in the use of public reason (as shown by Rawls’s use of supreme court as an ‘exemplar’ of public reason in *PL* 231-240), he never denies that citizens in general might have to discharge their ‘duty of civility’ (*PL*, 217) and address each other in terms of public reason. See for example *PL* 243: ‘what public reason asks is that *citizens* be able to explain their vote to one another in terms of a reasonable balance of public political values’ (and Rawls’s example of the discussion of abortion in this regard in the footnote on the same page). Or, similarly, (Rawls, 2005 [1997], pp. 444-445): ‘we say that ideally citizens are to think of themselves as if they were legislators and ask themselves what statutes, supported by what reasons satisfying the criterion of reciprocity, they would think most reasonable to enact’.

following Rawls, their advocates call 'reasonableness'. In Gillian Hadfield and Stephen Macedo's words, 'a shared system of public reasons thus constitutes a shared fund of common meanings that agents engaged in cooperation can express to one another to signal their mutual commitment to cooperation' (p. 46).

The question worth investigating is to see whether the use of public reason, as opposed to reasons derived from citizens' conceptions of the good, in the public realm, can indeed provide that type of assurance. Some problems have already been noticed. The first issue is whether the reliance on public reason in circumstances of public deliberation is, to use a Kantian phrase, a categorical moral imperative or is conditional on others' reciprocity (and therefore, in a Kantian sense, 'hypothetical'). Andrew Lister (Lister, 2017) has recently suggested the latter view, which I find plausible; however, Weithman himself in fact accepts it, as shown by the conditional nature of his presentation of *Political Liberalism's* 'Nash claim' ('at least when others live up to those values and ideals as well').⁹ Indeed, the very point of defending public reason as a possible solution to the assurance problem is to demonstrate that, aside for the moral reasons that impose public reason as a way to discharge the duties of legitimacy (for the state) or equal respect (for citizens),¹⁰ there are *pragmatic* reasons to rely on it publicly.

A more serious challenge comes from advocates of 'justificatory liberalism', who argue that public reason, understood *à la* Rawls as a public discourse that is significantly (although, in Weithman's interpretation, not absolutely)¹¹ devoid of

⁹ See also: 'justice as fairness will be "stable for the right reasons" only if everyone in the well-ordered society knows that everyone else is committed to living up to its values and ideals' (Weithman, 2016, p. 163).

¹⁰ Moral reasons that, I have tried to show in the previous chapters, do not really stand up to criticism.

¹¹ See Weithman's reference to the Rawlsian idea (expressed in *PL*, 247-254) that citizens are free to use their 'comprehensive' reasons publicly, provided they can then provide a further 'public' justification in case the exclusive reference to their own values were unconvincing (Weithman, 2016).

references to citizens' specific affiliations, only 'produces stability under unrealistic assumptions' (Thrasher & Vallier, 2015, p. 934). It is unrealistic because, in order for public reason to guarantee assurance, citizens need exposition to almost perfect information about their co-citizens' attitude regarding possible conflicts between the pursuit of their individual good and the compliance with public norms. Everyone needs to know that everyone else endorses a reasonable conception of the good that may overlap with other, equally reasonable, conceptions on a 'political conception of justice'. Absent that 'common knowledge', citizens would never be certain that their co-citizens can stop complying abruptly, even when 'complying' has been their ordinary conduct up to that moment.

But, how can that knowledge come about? The use of public reason cannot by itself generate knowledge, at least for the two concomitant reasons that it can be insincere and that is too easy. It is just too easy for citizens to couch their non-public concerns, derived from their sectarian conceptions of the good, in 'public' terms, so to better persuade their co-citizens of their good will. The expression of public reason would therefore amount, in the eyes of the critics, to a form of 'cheap talk', a technical term used in the game-theoretic literature to describe 'costless or very inexpensive, non-binding communication in a game' (Thrasher & Vallier, 2015, p. 943). Because of this, every signal sent by a citizen through the use of public reason risks generating its own interpretive disagreement. Citizens will be bound to wonder, for every display of public reason on the part of their co-citizens, whether that particular display was sincere, and therefore may constitute evidence of their reasonableness, or is just a case of 'cheap talk' that does not prove anything about the talker's attitude.¹²

¹² These two critiques (the one about the excessive ease of public reason and its possible insincerity) are referred to as 'noise' and 'amplification' in (Thrasher & Vallier, 2015) and (Kogelmann & Stich, 2016).

Some of the observations I raised in chapter 3 should show I am generally sympathetic with the critique, although with some important provisos. I do agree that the use of public reason can be insincere and even malicious. I even provided an example in that chapter about a debate that is clearly substantive and that some of the parties involved prefer to frame in purely ‘public’, non-sectarian terms so that they do not have to disclose their much more controversial ethical viewpoints.¹³ Following from the previous chapter, I would further point out that a scenario where citizens need to constantly question the sincerity of their compatriots’ public affirmations is one where equal respect is gravely compromised.

Nonetheless, we need not attribute to others a malicious attitude in their use of public reason in order to see problems with this model in terms of political stability. What I presented in Chapter 3 as the problem of interpretive disagreement suffices to show that political stability, if it is to depend on mutual assurance, cannot be based on public reason. To reiterate, the duty of civility demands that citizens refer to certain ‘political values’ when they defend their choice of voting on ‘fundamental questions’ (*PL*, 217). The political values in question are no other than the values agreed upon via the ‘overlapping consensus’ mechanism, or, in yet other words, the components of the ‘political conception’ of justice. But, when discussing a political matter in the public forum, citizens cannot just refer to those principles, because, as I previously showed, they are necessarily vague. Hence, they will need to produce instead a particular argument showing that the political principles and values agreed in the ‘overlapping consensus’ cohere, *under a specific interpretation*, with the norm that they are now proposing to enact.

¹³ See my reference to the Second Amendment debate in the US at p. 112.

Can their attempt at framing the matter in public terms, without referring to controversial judgements derived from their sectarian conceptions of the good, assure their fellow citizens that they are indeed willing to sacrifice the pursuit of that very conception of the good, in case the norm demands so? The first thing to notice here is that, if certain parties in the debate felt deeply involved in the matter, due to the 'sectarian' values they adhere to, and yet decided to defend their position purely relying on shared principles and values, then they indeed made a significant effort, so their talk is not that 'cheap' after all. However, their effort might be an indication of the most various things. For example, it can be an indication that the parties believe they have no other way to win the political debate unless they refer to values and principles that are at least minimally persuasive to others. The use of public reason in this sense is merely *strategic*. Alternatively, it can be an indication that, in the history of that particular society, certain types of considerations have been generally barred as unfit to a political debate.¹⁴ We can say that the use of public reason is here *contingent* on a particular history. Yet other explanations are, I am sure, available.

Because of this possibility of giving multiple explanations to the same phenomenon of public-reason display, I doubt public reason *per se* can provide any definitive assurance regarding the political attitude of public-reason adherents. The signal sent by public-reason users would require an interpretive effort to be decoded, and the interpretive effort can generate disagreements and possible mistakes. Even though we may expect that agents might improve their ability at interpreting correctly these signals coming from their fellow citizens' use of public rea-

¹⁴This is the situation, I believe, with purely religious considerations (say, biblical references) in most of Europe. See my example in footnote 10 in Chapter 3.

son, there is no certainty that stability will come about through this protracted exchange of variously interpretable information,¹⁵ or, at the very least, this is neither Rawls's thesis nor that of his followers.

There is, furthermore, an underlying normative problem. It remains unexplained why *I*, as an autonomous moral agent, should guarantee my co-citizens that I am not evading the normative requirements placed on me, despite my disagreeing with them, simply because such requirements have been defended appealing to principles I would normally consider acceptable. Again, it is worth reminding that my disagreement with fellow citizens who are defending a particular norm appealing to shared principles might be so deep that I eventually end up believing their interpretation is mistaken and possibly malicious. Should I still grant them my personal assurance that I am not going to evade the norm, purely because they have done me the favour of only referring to principles that were in principle sharable by me? That seems implausible, among other reasons because, had they really just referred to their own sectarian values (that are completely stranger to me), there would not have been any debate at all. The fact that some agents tried to discharge their duty of civility by initiating a debate does not seem to generate in other participants a duty of obedience with regard to the norms that are enacted at the end of the debate.

My scepticism about the use of public reason as an assurance mechanism is then double-pronged. On the one hand, I doubt the mere display of public reason may constitute reliable evidence that public-reason adherents are indeed willing to sacrifice their non-public values. On the other hand, I do not believe defenders of

¹⁵ This is the thesis, for example, of (Aumann & Hart, 2003): protracted 'cheap talk' can generate, through a process of gradual revelation of progressively more relevant information, a socially optimal outcome.

public reason demonstrate why citizens *should* provide that kind of normative assurance whenever they see that their fellow citizens, against which they are debating on some matter, make use of public reason.

The combination of these two points makes me agree, therefore, with those critics of Rawlsian liberalism who have captured the inherent ‘fragility’ of consensus-based models of stability.¹⁶ The problem is to see what a valid alternative is. For a ‘valid alternative’, I here mean a ‘liberal alternative’, that is, an alternative that does not exclusively rely on coercive means, of the type I described in the presentation of the Leviathan game.

The same critics of the Rawlsian model of stability observe that public reason is an unnecessary addendum in the quest for stability. The final aim of such quest could be seen instead as the emergence of a ‘choreographer’, that is, a ‘correlating device that implements a correlated equilibrium [...] in which all agents play strictly pure strategies’ (Gintis, 2010, p. 251). What Herbert Gintis refers to as a ‘correlated equilibrium’ is an equilibrium that obtains through the introduction of a further player, the choreographer itself. The classic example is traffic regulation. In the absence of traffic signals, drivers are required to find an equilibrium among themselves every time they reach an intersection, presumably through a costly bargaining process. Traffic regulation releases drivers from that particularly onerous burden.

The term ‘choreographer’ is misleading because, in practical terms, there is no difference between a choreographer and the Leviathan, apart from the fact that the authority of the choreographer is presumed to derive exclusively from her expertise, rather than from some value-arbitrary factors such as the possession of power. In both cases, the main players of the game (citizens in one case and dancers

¹⁶ (Gaus, 2011); (Thrasher & Vallier, 2015); (Kogelmann & Stich, 2016).

in the other) let the indications provided by the 'coordinator' (to use an umbrella term) pre-empt their own reasons for action in the domains where the 'coordinator' is active.

Hence, the problem is to see *how* the 'choreographer' can emerge in the circumstances of liberalism and *who* or what is entitled to play that role. To be fair to its advocates, we do not need to see the choreographer itself, despite the misleading name, as a completely external, authoritarian player who has the capacity to *impose* social norms on players that are incapable of finding an equilibrium by themselves, for example because they have very imperfect information about each others' attitudes. 'Choreography' can certainly emerge in all sorts of way,¹⁷ but only some of them will be compatible with respect for individuals' autonomy and justifiable, in liberal terms, to the 'choreographed' agents.

'Public reason' is, in the eyes of the Rawlsians, not so much a choreographer (because it is not a complete set of norms capable of generating a 'correlated equilibrium' by itself) as it is a liberal strategy to let choreography emerge out of individuals adopting a certain liberal vocabulary ('public reason'). The critique I previously advanced is that, because of the multiple interpretations we can give to the epistemic signals sent by other participants, using public reason to let a stable coordination system (a 'choreographer') emerge is both empirically dubious and morally questionable. A valid critique to the public reason solution to the stability/mutual assurance problem is therefore that it is an inefficient and morally questionable way to let 'choreography' emerge, and not that it is a strategy that disregards the role of choreographers in generating equilibria.¹⁸

¹⁷ For example, Elinor Ostrom's vastly praised research on the way in which various communities have managed common-pool resources without the need for an external intervention (in the form of state legislation) (Ostrom, 1990) can be interpreted as a reconstruction of the ways in which 'choreography' mechanisms can, in certain conditions, autonomously emerge.

¹⁸ Which is, by contrast, the critique that both (Gaus, 2011) and (Thrasher & Vallier, 2015) advance.

A real alternative to ‘public reason’, on the other hand, is using convergence-based strategies of persuasion in the deliberative process. To reiterate, under the convergence-model, public justification only obtains through the assent (or, more precisely, non-dissent) of all participants, who can claim that a public norm does not contrast with their (‘intelligible’, in Vallier’s model) individual reasons. Those who are trying, in the deliberative process, to defend a norm in a way that may achieve public justification, must provide reasons that can forestall ‘intelligible’ complaints. According to a recent contribution (Kogelmann & Stich, 2016), this effort in addressing other citizens’ ‘intelligible’ complaints is precisely the kind of assurance that is needed to guarantee stability.

The argument behind it is the following: as Rawlsian public reason is no other than ‘cheap talk’, because citizens can always couch their sectarian arguments in terms of public reason without incurring too many costs,¹⁹ those that intend to assure others of their willingness to comply with social norms need to engage in more costly activities. Constructing an argument that is intended to be persuasive for people adhering to views I do not personally adhere to is costly enough, but what is required here is *directly addressing* those views by constructing an argument that cannot be objected starting from the evaluative standards of the parties I am confronting.

That certainly shows some commitment. But, commitment to what? Brian Kogelmann and Stephen Stich, two of the proponents, seem to assume that, the more costly persuasion becomes, the more it indicates the persuader’s willingness to assure others of their willingness to follow social rules. But, in the previous paragraphs, I provided alternative explanations for the commitment and they do not

¹⁹ I disagree on this. See above.

seem to be ruled out once the costs of commitment increase. In a significantly diverse society, we can expect that people who want to persuade others will *need* to engage the sectarian reasons that these others possess. Whether this is done to demonstrate others the kind of assurance that is necessary for political stability, or for purely strategic reasons in the hope of eventually winning the debate, is still up to interpretation. 'Convergence' models of persuasion are therefore as subject as Rawlsian 'public reason' to the challenge of interpretive disagreement.

Regarding their moral assessment, on the other hand, 'convergence' models fare much worse than the Rawlsian 'public reason' alternative as they not only condone, but utterly demand, a patent violation of equal respect. They require indeed to engage in those kinds of substantive evaluations of others' values and conceptions of the good that in the previous chapter I claimed equal respect would require abstaining from. Such evaluations would be necessary to produce an argument that engages others' conceptions of the good, often with the explicit aim of disclosing their internal contradictions and idiosyncrasies. 'Convergence' reasoning leads therefore to those forms of manipulations of others' views that, I argued previously, equal respect would require to abstain from because they disrobe moral agents of that 'opacity' guaranteed by the attribution of 'as-if rationality'.

If a notion of equal respect makes sense at all, it can be minimal and extremely modest in its normative demands, but must not be subjectable to easy bargaining. There can never be any trade-off between equal respect and stability; in case a conception of stability is such that leads to compromising a normative demand derived from the value of equal respect, either the conception of stability itself or the conception of equal respect need to be revised. As I believe the conception of equal respect I proposed in the last chapter is minimal enough not to deserve any further relaxation, I suggest rejecting instead any conception of stability that may

conflict with it. As the 'convergence' model of political stability produces these results, its rejection becomes now imperative.

5.2 On Political Realism and Its Insight

I presented at the outset of the chapter a brief summary of the ideological stream in European thought known as 'political realism'. It emerged as an attempt to isolate politics as an independent field of study, distinct in particular from theological discourse; empirical political science in all its branches is its most direct offspring. There is no doubt, however, that its founders were what we would call today political theorists, as they did not refrain from advancing highly theoretical arguments about the interpretation of political reality and that they drew from them normative implications. The interesting question is to see whether the political theory of political realism may have a role in the solution of the problem I am now trying to solve, namely how to reconcile justice and stability. My conclusion in this section is that the re-emergence of political realism in the last decades does not succeed in carving out a theoretically interesting midway between so-called 'political moralism' and *Realpolitik*, but does provide us with a profound insight regarding the ambition and limits of our investigation.

The critique of 'political moralism' or the 'ethics first' approach²⁰ that Bernard Williams pioneered, but never managed to develop, in the latest phase of his career, includes in a prominent position the idea that disagreement about matters of value is a constitutive feature of politics. Because of radical disagreement pervading any political society, the 'first political question' can only be, 'in Hobbesian terms [...] the securing of order, protection, safety, trust, and the conditions of cooperation' (Williams, 2005, p. 3). But these Hobbesian premises do not stretch very far, for,

²⁰ This latter expression is from (Geuss, 2008).

Williams adds, 'it is a necessary condition of legitimacy (LEG) that the state solve the first question, but it does not follow it is a sufficient condition' (ibid.). The Hobbesian premises reduce even further when Williams acknowledges 'that there might be circumstances in which the only way to be LEG involved being liberal' (p. 4). Past tense notwithstanding, it seems plausible to assume that the circumstances are still obtaining. If they are, then the Hobbesian recognition that politics originates out of an original attempt to overcome conflict does not lead to a significant challenge to the classic liberal paradigm, the one according to which, to reiterate, a legitimate state is only one that can offer some sort of justification for its authority that is accessible to the citizenry at large.

The core objection of Williams seems therefore to be meta-ethical, if not semantic. He argues that the 'basic legitimation demand', 'which comes from a conception of what could count as answering a demand for justification of coercive power [...] is implicit in the very idea of a legitimate state, and so is inherent in any politics' (p. 8). Which implies, as he specifies later,²¹ that the loosely Kantian conception of autonomy and rationality, which, following a long liberal tradition, I have myself appealed to in the last chapter, can be at most one way to *answer* the legitimation demand but does not constitute the premise of the demand itself.

In historical terms, this is undoubtedly correct. Liberalism is hardly the first or the sole ideology used to justify the authority of the state to its subjects, as demonstrated by Hobbes himself or by theorists as diverse as Robert Filmer, Joseph de Maistre, or Carl Schmitt, all intent on demonstrating that the authority of the state does not in any way derive its legitimacy (if even such a term makes sense) from the consent of the governed.

²¹ 'One can invoke a liberal conception of the person in justifying features of the liberal state (they fit together), but one cannot go all the way down and start from the bottom' (pp. 8-9).

Historical contextualisation has a powerful humbling effect; it deprives ideologues of the illusion of the universal and a-temporal success of their theory. What could be the consequence of this ‘humbling’ spirit for the Rawlsian theory of legitimacy as public justification? On paper, public justification does not certainly sin of excessive ambitiousness.²² But, in practice, the conviction that there is a single liberal model of public justification that can solve the public legitimation demand, at least in the circumstances of a democratic society, is, Williams teaches us, destined to remain a delusion. A liberal model of legitimacy (such as Rawlsian ‘public justification’) might certainly ‘make sense’ (pp. 10-11) to most of us today, but ‘the conditions of LEG in modern states present a progressive project’ (p. 17)²³ that presumably is not going to be completed any soon. However, if the basic legitimacy demand must be met, and if the only way to meet it in the present circumstances is through some kind of liberal story, then there is simply no alternative to what Rawls and the other liberal theorists are doing, which is, to try and propose a ‘sensible’ model of LEG.

Instead of reconciling stability and disagreement, William’s contextualisation has led to scepticism about the feasibility itself of such reconciliation. Which is, in much less nuanced terms, the clear objective of those hard-core political realists, such as Raymond Geuss,²⁴ interested in emphasising the ‘conflictual’ or ‘agonistic’ essence of politics. For them, any account in political theory that aims at reducing,

²² Rawls notoriously claims that, ‘to formulate such a conception [political liberalism], we apply toleration to philosophy itself’ (Rawls, 1999 [1985], p. 388).

²³ For a Williams-based account that clearly expresses scepticism about the possibility of a full-blown theory of political legitimacy, see (Horton, 2012): ‘although there is a tendency to theorize political legitimacy simply as a property that a state either possesses or lacks, it also needs to be seen in terms of an ongoing process through which legitimacy is affirmed or denied’ (p. 144).

²⁴ (Geuss, 2001) (Geuss, 2008).

let alone overcoming, such conflictual dimensions, amounts to a pure 'displacement of politics' in political theory.²⁵

I find it difficult to see the point of this emphasis on the conflictual dimension of politics. I tend to believe it can be interpreted in two ways. One is, as I wrote before, a slightly less nuanced version of Williams's own 'humbling' point: reminding political theorists that the 'circumstances of politics', to follow Jeremy Waldron's phrase (Waldron, 1999, p. 102), include discord and potential conflict, enlightens them about the illusion of achieving stability through a universally successful account of legitimacy. This interpretation, *pace* its advocates, does not constitute a devastating challenge to the Rawlsian model of political theory. A Rawlsian can easily welcome the humbling spirit of the objection and recognise that public justification is an aspiration more than a universally fixed model, and as such perfectly fits within that 'progressive project' of legitimacy. Nothing within the Rawlsian account invites political theorists, to quote Williams again, to 'play Kant at the court of King Arthur',²⁶ nor, to stick to the metaphor, to play Rawls at the court of whatever sovereign (literal or figurative) happens to be reigning in the year 3000. Historical contextualism is not only compatible with the Rawlsian presentation of political liberalism, but somewhat underlies it, as Rawls makes explicit through his remark about political liberalism taking the lead from ideals that are 'implicit in our society' (Rawls, 1999 [1985], p. 396), and Charles Larmore more than any other clarified (Larmore, 1996, chapters 6 and 7).

The second interpretation denies any role to legitimacy; politics is *no other* than conflict, of vaguely concealed interests more often than values. I want to take

²⁵ The expression constitutes the title of a fortunate book by Bonnie Honig (Honig, 1993), and is variously cited in the realist literature (Galston, 2010, p. 386) (Baderin, 2014).

²⁶ As acknowledged in (Jubb, 2015).

this claim in its most extreme, Nietzschean, form, because I feel it somewhat underlies Raymond Geuss's claim that 'a political philosopher cannot take ideals, models for behaviour, or utopian conceptions at their own face value' (Geuss, 2008, p. 10). Of course, Geuss does not lead this remark to its extreme relativistic conclusions, although he does claim that 'modern politics is importantly about power, its acquisition, distribution, and use' (p. 96). Taking the Nietzschean insight to its extremes would lead not merely to the rejection of the question of political legitimacy (a conclusion Geuss, unlike Williams, is happy to embrace; see pp. 34-36), but to the complete loss of sense of normative political theory. If politics is merely a conflict of irreconcilable interests, then any attempt to somewhat contain such conflict will be no other than a more or less masqueraded attempt to let one of the interests in the field prevail.

Alas, we do not need much political theory to avoid conflict, even assuming that the end of conflict is either feasible (as Geuss seems to doubt) or even desirable (as a pure Nietzschean would certainly deny). What we need is an overarching moral meta-principle which says 'Pain is bad; do everything you can to prevent it', and all the rest (that is, the search for the best strategy to avoid the pain that is generated by conflict) is empirical. Hence, it seems that, if we follow the second interpretation of the emphasis of conflict in politics, we are led to either callous nihilism or to a form of crudely minimalist negative consequentialism in which one cannot even aspire to the maximisation of pleasure (for what counts as good is of course object of conflict), but to the minimisation of pain.

I do not have much to argue against this grim view of politics. My only comment is that it proudly ridicules the very objective that I set out in the present chapter, that is, the possibility of maintaining political stability without compromising the liberal spirit of acceptance of diversity. Authors who intend to take inspiration

from this tradition ought therefore to be wary of falling easily into the trap of reducing politics to conflict, because that would have the undesirable effect (for a liberal) to prevent any role for a minimally normative account of political legitimacy. Retaining a normative role for political legitimacy, it is worth emphasising, does not amount to claim that liberal political theory must include an account of political legitimacy that has universal validity and popularity. On the contrary, the considerations I raised in the three chapters I dedicated to legitimacy make me certainly agree with Matt Sleat that ‘the use of coercive power is [always] (at best) imperfectly legitimised’ (Sleat, 2014, p. 316).

But, if the objection against ‘political moralism’ is merely that no account of legitimacy will ever garner enough popularity in society to justify the exercise of coercion to every single citizens, it is based on an unreasonable requirement for political theory. The unreasonable requirement demands that political theory is capable of providing not merely plausible arguments, but also ones that can garner universal empirical acceptance. I see how this unreasonable requirement might be based on political liberalism’s account of legitimacy itself, as that does contain the idea that the legitimacy of a public authority depends on the *acceptability* of such authority to reasonably idealised agents. But acceptability and empirical acceptance are no synonyms.

Of course, for the supporter of the ‘grim view’ I sketched above, acceptability is meaningless, for the only thing that matters is the actual acceptance of the political order, demonstrated by citizens through their refraining from rebellion. Even for the extremely pessimistic consequentialist that believes the only moral maxim is the minimisation of sufferance, acceptability is irrelevant, for what matters, at most, is a justification of authority, however insincere and hypocritical, that may persuade citizens that refraining from taking arms against the government is in their interest.

Once we leave the 'grim view', in either its nihilistic or negative consequentialist presentation, the next step in the argumentation can only be the defence of a moral view (however minimal) that aspires at being acceptable; precisely, that aspires at being the *most acceptable*. Some of those who have refused to commit to the 'grim view' will try to play Kant at the court of King Arthur; they will conceive political theory as a quest for the eternally fixed moral truth. Others will accept Williams's central point that the responses to the most fundamental question of political theory are historically contextual. Others still can situate themselves on any intermediate point between Williams's historical contextualism and the belief in the a-historical validity of moral truth. What distinguishes all of these positions from the 'grim view', however, is the commitment to the idea that the aim of political theory is to produce justifications (for the authority of the state, in case we are interested in legitimacy) that appeal to something more than the might of the speaker ('might is right'), or a survival instinct that we presume is undeniably universal among humans (the negative consequentialist, loosely Hobbesian view). That 'something more', however minimal and vaguely specified, will constitute the backbone of a liberal account of political legitimacy.

It is time to draw out some conclusions from this digression into political realism. I certainly did not derive from that a solution to this chapter's main problem. In fact, political realists will probably question whether an investigation in political theory that aims at reconciling stability and disagreement makes sense at all. To them, this is a question that can only have context-specific, mostly empirical, answer. I say 'mostly empirical', for some of them, the ones unhappy to embrace the 'grim view', will accept that some normative requirements (legitimacy primarily) will constrain the range of strategies that the public authority can appeal to in its

quest of stability. But the choice of one among those normatively legitimate options will be an empirical matter.

The question of political stability, as I have so far presented it, is an investigation about the accounts of political legitimacy that, because of certain features, are more likely to be empirically accepted within a certain polity. The insight I derived from political realism is that this question, if addressed from within political theory, sits uncomfortably between the irrelevance and the infeasibility. It is both a question that lies outside the scope of political theory (or at least at its margins), and one that can never receive a fully satisfactory answer.

As I recognise this is an important insight, I need to reframe my question. I feel this objection, with the 'humbling' spirit it conveys, is particularly insidious against the accounts of political stability I considered, and already rejected, in the first section of the present chapter. Both the Rawlsian followers who defended the role of public reason as a pragmatic mechanism capable of generating mutual assurance of compliance and the critics who suggested a similar role for 'convergence' greatly relied on the expectation that political theory can indeed have that role. They relied, precisely, on the expectation that political theory can produce solutions to a practical problem, in the sense that it can indicate the use of which normative instrument is more capable of generating stability at least in the circumstances of a liberal state.

Moreover, although their reflections were (implicitly) contextualised for a broadly liberal society, their attention to context stopped there. Which means that the underlying supposition is that the same account of stability can lead to normative recommendations that are indeed successful across liberal societies, regardless of historical, economic, or political contingencies. To exemplify, what both Rawlsian and 'justificatory' liberals seem to be committed to is the idea that the same account

of stability, despite its strong relation with the theme of religious pluralism, applies both to states with a long tradition of religious disestablishment, such as the United States, and to states, such as the United Kingdom, where a certain form of religious establishment is the norm.²⁷ Or, similarly, it applies equally to both strongly multicultural and to homogeneous societies.

This confidence in the equal applicability of the same model of stability across liberal societies explains why the authors considered in the first section of this paper so often conceive the solution to the question of stability as the solution to an idealised game, the game of mutual assurance. But both the objections presented in the previous section and the insight I derived from political realism are sufficient to express my overall scepticism about grounding the whole question of stability in mutual-assurance game. As I showed in the previous section, mutual assurance cannot just depend on the common use of a single political tool, be it the exclusive use of public reasons in the public debate or the requirement of ‘convergence’ for public justification.

My previous point about the difficulty of interpreting a presumed signal of assurance can now be reframed following the realist insight. Interpretations, realism wisely reminds us, are based on a particular interpretive context, made up of a plurality of contingencies, relative to the interpreter’s social, economic, and cultural status (at the very least). Abstracting from these contingencies is perfectly possible, but one should wonder whether it is, after all, useful, when the objective is as concrete as producing an account capable of guaranteeing political stability across liberal democracies. In other words, there is, in the accounts I considered in the first

²⁷ On this point I was inspired by (Raz, 1998, p. 52), which, I believe, preceded political realists in emphasising the role of contingency in achieving political stability.

section, a mismatch between the absolute concreteness of the objectives of the investigations and the abstract, context-independent methodology of the investigation itself.

The insight of political realism consists, as I see it, of a memento regarding the ambitions and limits of the 'liberal project'. As I already remarked in the methodological section of the Introduction, ignoring or abstracting from context in normative political theory has its consequences, the first being that, the less context we take into account, the less a particular account will be capable of generating immediately action-guiding normative recommendations for the here and now. The conclusions new-wave political realists derive from this is that political theory ought to abandon any universalistic ambition and focus instead on the context under which certain normative accounts become applicable and useful. The conclusion I prefer to draw, on the other hand, is that political theory ought not to give up on its universalistic ambitions, but must pursue objectives that are aligned with the abstraction from contingencies that characterises its methodology.

So, when political theory addresses the question of political stability, the extent and applicability of the account of stability it can produce will vary depending on the contextual data that are taken into account. Highly contextualised inputs will generate immediately applicable normative outputs. Which means that, by choosing to retain the universalistic aspirations of normative political theory, I also accept its limited role, when it comes to the elaboration of immediately applicable action-guiding norms. It follows that, insofar as I am considering the question of political stability, the outcome I expect from such investigation cannot be the justification of an account capable of setting out all the necessary and sufficient empirical conditions under which political stability comes about. This type of overly optimistic expectation, which I believe underlies the accounts of political stability I considered in the

previous section, is hopeless, unless robust contextual data are factored in. Which will imply, however, that idealisations of the form that we see in those works, leading authors to conceive stability as a game, are no longer permissible. In the next and final section in this chapter, I aim to present my own take on political stability, which hopefully lies successfully between the two extremes of absolute idealisation and absolute contextualism.

5.3 Stability-Enhancing Toleration

5.3.1 From Rawls's 'Aristotelian Principle' to Toleration as Equal Entitlement

I expressed in the first section my admiration for the idea, that we retrieve originally in *A Theory of Justice*, that stability must depend on the demonstration of the congruence between one's conception of how to conduct a good life and one's adherence to the norms of society. In order to demonstrate that, one can focus on either of the two terms of the equation. Rawls in *Theory* tried to demonstrate that the 'sense of justice' that people gradually acquire by living in a well-ordered society is consistent with a 'rational plan of life'. Later on, after what I defined before as his Copernican Revolution, Rawls decided to focus instead on the conception-of-the-good end of the equation, showing that reasonable conceptions of the good can overlap on a political conception of justice that reasonable citizens will accept as a plausible justification for the authority of the state.

Before I present my account, it is worth considering how Rawls believed the 'congruence' between the rational good of any individual and the social right could be demonstrated. Rawls's solution is grounded in a theory of the good that we may call 'minimal naturalistic perfectionism'. I define it this way because of the inclusion, within it, of a principle, labelled 'Aristotelian', used to discriminate between superior and inferior ends that people might pursue. The principle runs as follows:

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 complexities*. The intuitive idea is that human beings take more pleasure in doing something as they become 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*. (TOJ, p. 374. My emphasis.)

The inclusion of the Aristotelian Principle is so much interesting, considering that in working out his theory of ‘goodness as rationality’, Rawls had in the previous sections mostly relied on a framework that, although not collapsing onto a purely instrumentalist, Humean model, was significantly neutral regarding the selection of ends that ought to inform a rational plan of life.²⁸ The Aristotelian Principle comes to correct this end-neutral approach, providing an important metrics to gauge the worth of capacities that people might intend to develop: the more complex the capacity, the greatest the satisfaction in achieving it. Hence, even though certain capacities are necessarily too complex to generate the satisfaction of achievement during the limited time of a lifespan (TOJ, p. 376), it is generally rational to pursue, among achievable ends, the ones that contain ‘more intricate and subtle discriminations’.

²⁸ I am here referring to sections 60-64. Whereas in the previous sections Rawls had mostly relied on economic literature, where instrumental rationality is the norm, in presenting the Aristotelian Principle he expresses his debt to Philippa Foot’s scepticism about the reduction of goodness to a matter of choice (Foot, 2002).

The Aristotelian Principle serves the fundamental purpose of demonstrating that the pursuit of the sense of justice, which is ‘an effective desire to apply and act from the principles of justice and so from the point of view of justice’ (TOJ, p. 497) is, within a ‘well-ordered society’,²⁹ a rational pursuit. Rawls almost concludes *A Theory of Justice* by pointing out that ‘it follows from the Aristotelian Principle [...] that participating in the life of a well-ordered society is a great good’ (p. 500). The argument is fully perfectionist in nature and re-elaborates an old insight, that we usually associate, at least in the context of modern-age philosophy, with Marxian ideas.³⁰ Only within a well-ordered society, Rawls argues, can the various form of human activity be fully realised, for ‘given the social nature of humankind, the fact that our potentialities and inclinations far 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’ (ibid.).

The Aristotelian Principle, in sum, with its imperative to pursue those activities that contain greater complexities, imposes on moral agents that they abide by the rules of a society that better enhances their flourishing. As complex activities require cooperation, and cooperation requires rules of conduct, moral agents are happy to abide by those rules of conduct that (1) they themselves have agreed upon under the fairest conditions (TOJ, p. 501) and (2) make possible the realisation of those activities that necessitate higher levels of coordination and that are inherent to the ‘social nature of mankind’. Of course, it is perfectly possible to refuse this defence of ‘congruence’ as resting on an implausible and overly optimistic picture of

²⁹ TOJ, chapter 8 had tried to demonstrate the emergence of the ‘spirit’ in the circumstances of a well-ordered society. In particular, Rawls mentions three psychological laws (§75) that are supposed to create ties of affection between each citizen and the public institution she is living under.

³⁰ For a 21st century’s re-elaboration, see (Cohen, 2009).

human nature; the holders of the 'grim view' will certainly sneer at the mere suggestion. But, against those, Rawls concludes, we can claim that 'their nature is their misfortune' (p. 504).

Now, as I argued at the outset of this chapter, my intention is somewhat to defend the Rawls of *Theory* against his older self, in the sense that my account is going to be one where stability depends on a particular moral attitude that citizens need to cultivate. My argument does not depend however on the type of naturalistic perfectionism that underlies the Aristotelian Principle. This not because I reject the principle, or I believe it is an inappropriate basis for showing the congruence between the right and the good. Rather, I prefer to focus on aspects that are more directly relevant to the question of stability where this is understood, as in the original formulation, as the preservation of a just political establishment.

The moral attitude I intend to bring to the fore is not a component in human excellence, or at least I do not intend to defend it that way. I defend the attitude, by contrast, for its consequential value, for its being instrumental to the possibility itself of there being, in the context of a liberal society, political stability.

I define the attitude 'toleration' to follow a particularly fortunate formulation provided by Thomas Scanlon. However, it should be clear that my intent is not that of defending it *as the best account of toleration*;³¹ I simply define it this way to follow Scanlon's terminology and I will compare it to other models of toleration only when these too are concerned with stability.

Enough with preliminary remarks. The attitude I refer to is one in which citizens attribute to each other a particular moral entitlement, namely, the moral entitlement to try to shape society according to criteria they would consider reasonable.

³¹ My account does certainly not amount, therefore, to a conceptual analysis of toleration, as the one pursued in (Cohen, 2004).

Notice that I am here using 'reasonable' in the ordinary-language sense of the expression, the one that refers to the limits one believes are fit for the imposition of rules on others, once others' needs and interests are (roughly) taken into account. Thus, the moral attitude I am here focusing on does not denote passive acceptance of others' imposition of rules. By being tolerant, I do not simply give those of my fellow citizens who happen to compose the majority carte blanche when it comes to the selection of my society's public norms. But I do accept that what my fellow citizens unilaterally call 'reasonable' norms might be ones I would not like to see prevailing in my society.

In a way, I adopt an attitude that is much more demanding than passive acquiescence to the will of the majority (although, I would add, much less irresponsible), as it requires that I positively represent my fellow citizens as individuals capable of formulating rules that incorporate some consideration for my own interests and needs. Only once that representation of my fellow citizens is there, I can accept that, despite my dislike for some of the policies my fellow citizens are trying to impose, I respect their entitlement to do so. Indeed, the value of toleration, to quote another of its (sceptical) advocates, 'lies partly in its difficulty, in its requirement that one should rise not only above one's own desires but above one's desire to secure the fullest expression of one's own values' (Williams, 1998, p. 26).

5.3.2 The Rejection of Modus Vivendi

Other than being more demanding than mere acquiescence, the tolerant attitude I have in mind is also significantly more demanding than the mere *modus vivendi* toleration that can be adopted by the least nihilistic advocates of the 'grim view'. Their view of tolerance is really at most a 'second-best option', which we adopt simply because it is 'a vast improvement over the sectarian bloodshed that we hear of every

day, in many parts of the globe' (Scanlon, 2003 c, p. 187). Nothing is wrong, of course, in this reasoning, and the humbling spirit of political realism should remind us of the importance of this type of attitude in the resolution of conflicts. However, there is something unsatisfactory behind the grounding of stability purely on the avoidance of pain.

The lack of satisfaction is not necessarily practical, because I assume that the avoidance of the pain generated by enduring conflict can be an extremely compelling motive in the gradual achievement of political stability. Furthermore, *modus vivendi* toleration can be defended in more sophisticated ways than the ones the supporter of the 'grim view' would produce. For example, Chandran Kukathas grounds his defence of *modus vivendi* toleration in an epistemological argument, centred on the 'Kantian view that upholding reason requires a realm of freedom protected by toleration' (Kukathas, 2003, p. 131). In the absence of a reliable authority with 'privileged access to truth' (ibid.), only complete and absolute freedom of conscience, which takes the form of a *modus vivendi* toleration of *all* the views that are expressed in society, can provide the ideal terrain where the progress and development of practical reason can take place.

Kukathas, inspired by Elinor Ostrom's studies on the self-management of 'common-pool resources'³² is confident that, in conditions of toleration, a 'convergence of moral practices' (p. 134) will come to constitute the public sphere, and stability will therefore come about without any form of authoritarian intervention. Various communities of values will co-exist under rules of mutual toleration they have converged upon, without the need of a public morality or the sharing of some basic common values across the various communities.

³² See footnote 17 above.

Kukathas's account of toleration combines two different arguments. One is the classic Enlightenment view (which he reconnects to Kant but that we may also find, for example, in Mill's *On Liberty*)³³ that, once we abandon the view that truth can only be apprehended through some esoteric contact with an epistemic authority, only the most widespread freedom of conscience and the free circulation of ideas that originates out of it can guarantee advancement in human knowledge. The other is the libertarian view according to which the state, when trying to instil a (however minimal) public morality, is engaging in an activity that is both unnecessary and highly dangerous; citizens of freely adhered communities should instead be left free to negotiate the appropriate terms of cooperation that apply to them. The two views are compatible, but one can easily buy one and not the others. In fact, religious communities will presumably reject the Enlightenment view with its scepticism about epistemic authorities, but, as long as they refrain from coercing others into conversion and accept to be part of a negotiation effort to regulate society, they can easily welcome the spirit of the libertarian view. It follows that the two views play a much different role in the argument; the first, being controversial in a pluralist environment, can be at most persuasive for some of Kukathas's academic readers, whereas only the second can feasibly play some role in the achievement of political stability. Hence, even for Kukathas, what ultimately justifies the recourse to the particular policy of toleration he defends is the expectation that *modus vivendi* toleration will bring about stability in a way that is both more efficient and less risky than any attempt to condition stability on some minimally shared set of values.

What justifies, therefore, my previously mentioned lack of satisfaction for a pure *modus vivendi* model of toleration and stability? One source of scepticism is borne out of the previous reflection on the 'realist insight'. If stability is portrayed

³³ (Mill, [1859] 2015).

as either the avoidance of conflict or an exercise of negotiations, then political theory that works on stability is significantly far from its comfort zone. Other than pointing out some successful examples of cooperation that did not require the presence of a shared public morality, political theorists cannot contribute much to a discussion that is ultimately empirical as it concerns the social and political conditions that may, more than any other, help preserve a (possibly virtuous) political unit.

The deeper source of scepticism lies, however, in the definition itself of stability, which, in my view, cannot be a label we indifferently associate to either just or unjust political environments. The stability of an unjust political system is not, by itself, a moral value, even though some stable and unjust regimes might guarantee better living conditions to their subjects than some extremely unstable democracies. Even more importantly, the stability of a just regime is not a moral value either if it is brought about by significantly compromising the justness of the regime itself. Again, it might well be that certain less-than-perfectly-just but stable political regimes can guarantee better living conditions to their citizens than states that are perfectly just but terribly unstable, but this is another question that we will only be able to answer once we take into account the contextual data of everyday politics.

Moreover, I am unpersuaded by the view that a liberal state that tries to achieve stability should only be concerned with the *modus vivendi* co-existence of its citizens and should refrain from the promotion of a (however minimal) public morality that may help solidify the justice that the regime is supposed to implement. This restriction can be justified either empirically or morally. Empirically, the idea that liberal states are incapable of promoting public values is exaggerated. Liberal regimes will luckily be unable to exploit the most dramatically coercive means to promote the spread of ideas that they favour, but that does not make them helpless

in that regard. Both fully coercive and softly paternalistic, nudge-like norms³⁴ will shape the public morality prevalent in society; even more importantly, very few liberals would argue that the liberal state ought to refrain from managing or at the very least regulating the education sector, which has of course the most magisterial role to play in both the diffusion and containment of values and ideas.

Which immediately leads to the moral question of what liberal states are allowed to do in their attempt at seeing that certain values or ideas, such as toleration or equal respect, are widespread through the polity. Unfortunately, doing justice to this question alone would probably require a philosophical work as extended as the present one; hence, I need to make some assumptions, which, however, are going to play only a minimal role in the argumentation. I will assume that a liberal state is reasonably in control of the education system and that the education system itself can be organised so to discourage both current students and, by influence, their next of kin, to adopt blatantly intolerant conceptions of the good.³⁵ I have in mind such educational policies as memorial events, promotion of intercultural and inter-social dialogue, and, more generally, the requirement that schools (both public and private) should be organised so as to welcome the most diverse possible intake of students. Following my prior rejection of the ‘recognition respect’ model, I am much less inclined to accept, on the other hand, that the state may adopt policies that directly discriminate intolerant conceptions of the good, say by refusing to accord them the recognition (in terms, for example, of warranty on their prosperity) it already accords to other conceptions.

³⁴ Such as the ones presented in the ‘libertarian paternalistic’ model famously defended by Cass Sunstein and Richard Thaler (Sunstein & Thaler, 2003).

³⁵ Jonathan Quong, in the final chapter of his (Quong, 2011), tries to defend the thesis that this ‘containment’ of unreasonable doctrines is not in contrast with the strongly anti-perfectionist version of political liberalism he endorses. Luckily, as I am not myself endorsing full-blown anti-perfectionism, I do not have to defend these assumptions within a globally anti-perfectionist framework.

As I wrote before, however, these assumptions do not play a fundamental role in the argumentation that will follow. In the next subsection, which will conclude the present chapter, I am going to offer a more detailed characterisation of what I mean by toleration and why I believe the diffusion of a tolerant attitude in society is instrumental to the achievement of a stably just political order. I am further going to show how the justness of a political system renders toleration a somewhat less demanding option. I am not offering, therefore, a precise examination of toleration-enhancing policies. The few examples I provided in the last paragraph can be combined with the ones that the reader will find in the following chapter. Their role is simply that of showing that enhancing toleration within a liberal regime is not only possible but also necessary for the survival of the regime itself.

5.3.3 Toleration as a Virtue

I suggest focusing now on the attitude of toleration as something that people would *need* to gradually adopt, in their role of citizens of a fair, liberal, society; indeed, as a (if not *the*) liberal virtue. In being a virtue, it cannot be confused with the mere ‘putting up’ attitude that is the most to be expected from somebody embracing the ‘grim view’. Firstly, ‘putting up’ with something that is previously disliked, although sometimes commendable for its consequential benefits, can often display the possession of all sorts of malign prejudices, as with a father ‘putting up’ with his child’s homosexuality. Secondly, even when toleration as ‘putting up’ does not disclose the possession of prejudices, it might indicate instead a mere concern for one’s self-preservation, which can hardly constitute a virtue. In other words, the virtue of toleration cannot encompass such things as tolerating what ought not be tolerated (in the double sense that it ought to be respected as such rather than ‘put up with’ or that it

ought not to receive any form of respect at all)³⁶ or tolerating when the alternative of not tolerating is significantly risky for the tolerant party. Again, as in my previous response to Williams, whether the virtue of toleration historically originated out of these far from virtuous attitudes is a question that I take to be irrelevant to the justification of toleration itself.

In defending toleration as a virtue, Scanlon mentions in a footnote a famous statement that is often mistakenly attributed to Voltaire, and that, not to bore the reader, I will not directly quote here.³⁷ Constantly heard statements usually sound tacky, but this contains more than a grain of truth. It expresses the spirit that ought to inform citizens who are committed to seeing each other as equally entitled to shape society according to what they consider reasonable criteria. Indeed, if I intend to allow my fellow citizens to have a say in how the society in which I live can look like, I need to be well-disposed towards the ideas they might bring to the public arena; not in the sense that I need to like them, or that I need to refrain from fighting them if I believe they are wrong, but in the sense I need to accept those are ideas they are entitled as I am to propose as grounds on which to base our society.

This positive disposition towards others reconnects to the conclusions I raised in the chapter on equal respect, in particular concerning the role of 'trust'. I had concluded there, inspired by Jonathan Wolff's reflections, that a commitment to equal respect entails a form of trust, although the trust that is here required differs from the one necessary to attribute 'as-if rationality'. Relying again on Rawlsian categories, we might say that, whereas 'as-if rationality' respect entailed the assumption of others' possession of the capacity to form and revise a conception of the good, the virtue of toleration requires by contrast the assumption in others of the other

³⁶ For a more articulated discussion of these two limits on the conceptualisation of toleration as a virtue, see in particular (Horton, 1998).

³⁷ Scanlon mentions it at (Scanlon, 2003 c, p. 197).

Rawlsian 'moral power', the sense of justice. The object of trust is, in this context, the ability of others of acting as reasonable legislators, in the sense that their imposition of rules in societies will incorporate some form of consideration for their fellow citizens' needs and interests.

In other words still, the virtue of toleration centrally consists in the realisation that the 'grim view', despite its intuitive strength, is a representation of human nature that is worth rejecting. Again in a manner similar to the argumentative style I adopted in the chapter on equal respect, the reasoning is here characteristically circular, in the sense that I do not claim the rejection of the 'grim view' depends on the truth of some higher-order principle. On the contrary, I defend the rejection of the 'grim view' as an attempt to overcome a stalemate; as long as the 'grim view' remains the dominating paradigm, what Scanlon calls the 'informal politics' (p.p. 190 ff.) of a given society, that is, the sum of attitudes that citizens display in the public arena, especially with respect to those who disagree with them, will be characterised by a mutual distrust that will force citizens into costly bargaining procedures.

At the same time, as long as the 'grim view' is given some credit, many of the most central features of liberalism will keep resting on somewhat unstable ground. Take first the case of freedom of expression, which Scanlon himself mentions (p. 189). Any attempt at defending freedom of expression without committing to some form of principled (and not merely 'putting up'-type) toleration seems hopeless. We can try to defend freedom of expression in a purely prudential way, as a fundamental need that we want to be guaranteed firstly for us, and that only grudgingly we allow others to possess. But this defence of freedom of expression is peculiarly limited and unstable; it is limited to the types of expression every agent would consider most

central to herself and is subjected to the infinite give-and-take of political bargaining. Alternatively, we can defend freedom of expression in a completely impersonal way, by pointing out the advantages of a culture of freedom of expression for the achievement of truth. But this conception is also limited; even though we cannot predict which types of debates are going to be more conducive to the advancement of truth, we have strong intuitions that, say, scientific freedom is in this regard more relevant than freedom in the exercise of satire.

The commitment to freedom of expression must therefore have a different foundation, one that does not focus principally on the positive benefits of free expression itself for either the self-interested subject or some external value such as truth. More specifically, any kind of justification of freedom of expression that does not include the prior acceptance that others' ideas are as entitled to enter the public discourse as mine are, *whatever the consequences*, and that therefore these consequences are not going to be terribly dire, seems to produce only a very unstable defence of the ideal.

Similarly so with democracy itself, and more specifically with the practice of equal participation. We may appeal to all sorts of theorems to demonstrate the wisdom of the many;³⁸ let alone the fact that such theorems are usually countered by evidence leading in the opposite direction,³⁹ I doubt they persuaded anyone horrified by certain electoral results about the value of democratic procedures. And, since we are here concerned with the stability of the liberal order, it seems that an unconvincing justification of democratic equality provides a very fickle base on which to construe a liberal democratic order.

³⁸ (May, 1952); (List & Goodin, 2001); (Goodin & List, 2006).

³⁹ (Brennan, 2001); (List & Pettit, 2002); (Kornhauser & Sager, 2004). For a comprehensive refutation of the various arguments that aim at justifying democracy instrumentally, see (Kolodny, 2014).

The virtue of toleration, when applied to the case of democracy, demonstrates probably its best potential. The trust it demands is precisely what we need to justify equal democratic participation. The best justification for the inclusion of an element of trust within a minimally persuasive defence of democracy is again counter-factual. Absent any element of trust for other' ability as legislators, including their consideration for fellow citizens' interests and needs when proposing public norms, the justification of democratic procedures is at best partial. A merely partial justification will have a limited role to play in the achievement of stability; it can undoubtedly constitute a good ground for embracing democratic procedures, but only insofar as one can demonstrate immediate advantages that democracy generates for every citizen. But this is an almost impossibly demanding task, as the advantages that each citizen derives from living within democratic institutions are often varied and somewhat nebulous; capturing them under a rubric that is presumed to be universally persuasive is going to involve an infinite series of investigations.

Furthermore, it might make more sense to let every citizen free to identify the personal advantage that each of them finds particularly persuasive within the democratic ideal. If empirical stability is what matters, then that looks a more promising strategy, as in the end every citizen, having worked out, out of their own personal criteria, what is the best they can derive from their living under democratic institutions, will be perfectly motivated to keep supporting them. But, I argued at the end of the previous section, producing a 'convergence' strategy of this kind is neither a feasible nor a desirable aim for normative political theory. Given the impracticality of that task, normative political theory can still do a great service to the democratic cause by describing the attitude that would ideally need to prevail in the polity to increase the chances that the democratic order is going to remain stable.

I have so far mostly tried to demonstrate the relevance of toleration for political stability as the appropriate attitude that citizens ought to adopt in order for a just democratic state to be stable. It remains to be shown, however, that the virtue of toleration is one that citizens can adopt in circumstances that are not too idealised. It might be indeed that I have so far relied on an overly optimistic picture of human nature that is not less detrimental to the elaboration of a good account of stability than the 'grim view' itself. In case it turned out that the virtue of toleration were one so demanding for the citizens that it actually deprived them of the possibility itself of having a flourishing life, or that it significantly compromised their autonomous choice of a life worth living, the reliance on it to address the question of stability would be terribly inconvenient.

There is no point in denying that toleration is a demanding attitude. It is demanding in its premises, as attributing to fellow citizens a capacity to propose fair and reasonable terms of cooperation might often run against our best common-sense judgement. Similarly, it is demanding in its actual expression, as it requires moral agents to commit to the idea that their co-citizens, with which they can deeply disagree, are as entitled as they themselves are to have a say in the definition of society. Furthermore, if the good disposition towards others is turned into an expectation that others will indeed adopt a public behaviour that, despite not completely agreeable, one might find honourable and worthy of respect, it is an expectation that is going to be constantly frustrated. In this regard, the supporters of the 'grim view' will often have evidence on their side.

There are, however, some factors that characterise liberal societies that might render toleration a somewhat less inconvenient option. Toleration would be a terribly self-frustrating attitude, for example, within a polity where one specific minority of citizens never gets to have a significant chance at influencing the way in

which the society is structured, and whose interests are never taken into account in public decisions. A person belonging to that group would feel no reason to adopt a tolerant attitude towards her fellow citizens who maintain her and her group in that condition of decisional powerlessness. Hence, in order for toleration to be a convenient option for everyone in society, we need to assume that the society in question is one in which, to follow Waldron's terminology, the 'decisional majority', that is, the set of people that wins the debate and has the final say in a particular political decision, is not constant in time.⁴⁰

A functioning liberal democracy, however, is one where such examples of complete and irrevocable 'tyranny of the majority' are not supposed to take place. Now, one can say I have here provided a very good argument for my imaginary opponent, for I cannot deny that present democracies do include examples of what I called before 'decisional powerlessness'; examples, that is, of minorities whose ability at influencing the political agenda and having their interests publicly taken into account is regularly impaired. But what my argument implies is that the society in question is not going to be a democratically stable one, precisely because members of the oppressed minorities will justifiably feel no reason to adopt a tolerant attitude towards a majority that has shown in their regard only fear and contempt.

Which in turn shows that there is an assurance question affecting my account of stability too, but its solution is much simpler than in the previously analysed models. Whereas in stability modelled on public reason the correct understanding of signals sent by citizens publicly through their use of public reason would have required, I argued above, a considerable interpretive effort, identifying a signal of toleration, or a lack thereof, is much simpler. Of course, mistakes are still possible, but there are unequivocal cases. A majority that refuses any form of accommodation to

⁴⁰ (Waldron, 2006, p. 1396).

the requests of a minority, conscious that this will perpetuate the frustration that the minority already feels, will demonstrate an attitude that is the opposite of toleration. By contrast, the recourse to compromises and the acceptance of accommodations even when those are not strictly required by reasons of impasse – in cases, that is, where there is a majority that could unilaterally decide – gives the minority the strongest signal that the rest of the population does take their view into account. Like all other virtues, toleration has in this sense a capacity to generate a virtuous circle, in the sense that its public display by some can set a positive precedent that it is plausible to assume will be followed by others.

What about the frustration of those, however, whose decision to adopt a tolerant attitude has led to a disappointing frustration of all the expectations that had been placed on others? The best response, I believe, is the one Rawls gives in one of the very last pages of *A Theory of Justice*, that this question is, in the end, 'on a par with the hazards of love, [...] and the loves that may hurt the least are not the best loves' (TOJ, p. 502). The trust in others that, I argued before, necessarily accompanies and precedes toleration exposes the tolerant to a vulnerability that can lead to all sorts of misfortunes. But, as in love the vulnerability of one may sometimes be the strongest demonstration of the depth of the feeling, and thus lead to shared happiness, so in politics the vulnerability of a party may sometimes be needed to achieve shared ends. As in love so in politics, hence, commitment to what is valuable sometimes requires unilateral sacrifices.

5.4 Final Reflections on Justice and Stability

This final defence of the spirit of toleration might appear a romantic idealisation of politics, as it certainly is, up to a point. But I do believe that emphasising the idealistic elements of politics is as important as studying the details of its often far from

ideal reality. As I recognised before, political realism plays an insightful role in reminding political theorists of the limits of ideals, especially when it comes to the elaboration of a theoretical account that is presumed to have a direct practical function, such as, in the context of this chapter, guaranteeing stability. But stability would not even be a question, or at least not one minimally relevant to political theorists, were it not for the prevalence within a polity of certain ideals, however minimal, that I tried to describe in this chapter. It would not be a question relevant to political theory, because, in the absence of a shared set of reasons motivating each citizen to accept their fellows' entitlement to 'contribute to the definition' of their society, ensuring that each citizen will stand by the societal norms even when those go against their personal interests would be a matter of pure coercion.

There is, in fact, an alternative to absolute coercion, and is absolute justice. Were the state perfectly just, all citizens would have moral reasons to respect its norms. In a sense, this would be a solution to the question of stability, although a purely ideal one, because, in order for that to be functioning, we need to assume not only a perfectly just state, but also that citizens can be in such an epistemic condition as to all recognise the state as just. Notice that this type of idealisation could lead to direct practical implications – only, of a very pessimistic kind. It would entail that stability can only occur in the purely ideal circumstances of absolute justice and absolute knowledge about justice.

In the present chapter, I have tried to demonstrate that there is a third way through which stability can be said to prevail between these two extremes of absolute coercion or absolute justice. The presentation of toleration as an attitude capable of enhancing political stability follows this narrow path between two equally helpless ideals. However, in describing toleration as 'enhancing', as opposed to ensuring, stability, I hope I made sufficiently clear what are the limits of my account.

The prevalence of a tolerant attitude in society does not constitute either the necessary or the sufficient condition for the actual occurrence of stability. It constitutes, on the other hand, the necessary and sufficient condition for political theory's *interest* in the theme of stability.

Let me conclude with some final notes regarding the relationship between justice, stability, and disagreement. Stability, as I have presented in this chapter, is neither a purely 'external' constraint on the achievement of justice nor a value that we can purely derive from justice itself. It is, if possible, a combination of the two, in the sense that stability includes both elements that cannot be justified by purely appealing to moral reasons and elements that are, by contrast, strongly moral.

Among the first elements are those infinite contextual contingencies that, the realist insight reminds us, are massively relevant in the achievement of actual stability. To exemplify, let's take the United States and the United Kingdom as both countries that have equally succeeded in reaching a certain level of stability in the face of religious disagreement. What is remarkable is that an altogether similar level of (imperfect but somewhat functioning) stability is brought about through the opposite public treatment of religion: constitutional religious disestablishment but vast use of religious reasons in the public arena on the one hand, *de jure* religious establishment but *de facto* scarce reliance on religious reasons in the public arena on the other. How such a similar level of stability has been brought about, despite such different public treatment of religious pluralism, is a question that will necessarily elude any investigation in political theory, because very plausibly depending on far more contextual data than political theory can muster.

But political theory can refocus the attention on what most plausibly these cases have in common, which is a shared interest, within the polity, in remaining united even in the face of deep and irresolvable disagreement. And it can provide a

further and more articulated description of this shared interest, which is what I attempted to do in this chapter.

To refer back to a previous metaphor, the model of stability I have presented does not therefore correspond to either a perfectly choreographed ballet, where the choreographer has absolute discretion in the direction of every single move in the dancing routine, nor to some sort of leaderless dancing commune, where every move has to be negotiated among the dancers. It corresponds, by contrast, to a valued practice,⁴¹ where each of the agents involved is happy to sacrifice part of their individual interests, not necessarily in exchange for a future gain, nor because they are coerced to do so, but because they deeply care about the continuation of the practice itself, and in particular about the continuation of the special relationship they entertain with the other members.

Even though we can well imagine circumstances where that ‘special relationship’ is not one of equality (maybe the slavers did care deeply about their ‘special relationship’ with the slaves, although I doubt they could have found a way to justify that), in the (fortunate) circumstances of a liberal democratic society, that relationship must incorporate an element of equality, which I have tried to capture in the present and previous chapter. Equal respect of others as holders of intelligible views and acceptance of their equal entitlement to try to shape society according to those very same intelligible criteria are therefore not only moral demands, but also fundamental prerequisites to a flourishing and stable communal experience under conditions of justice.

⁴¹ Rawls’s analogy between (good) citizenship and love can be stretched further if one takes love, following Niko Kolodny, as being itself a ‘valued practice’ (Kolodny, 2003).

6. The Possibility of Accommodation and the Rejection of the Integrity Paradigm

The last chapter concluded with a rather abstract defence of the spirit of toleration that, I argued, ought to characterise the public morality of a stable liberal society, which followed an equally abstract defence of an account of equal respect that should inform the public treatment of differences in a pluralist environment. I feel I now owe the reader a more concrete presentation of the political response to the problem of pluralism that follows from those abstract liberal principles. The present chapter will do that by considering a string of policies that have received particular attention in recent scholarship in legal and political theory: accommodations for conscientious reasons.

The discussion of accommodations in this chapter is not intended as an exhaustive treatment of the political response to pluralism my previous reflections would recommend. It is to be seen, instead, as an example of the way in which my account can successfully address policies that are already there and that are the object of critical attention in legal and political theory.

More specifically, my aim is twofold. I will show first that a prominent paradigm that the philosophical literature has relied on to justify accommodations – the idea that accommodations are justified as a way to protect citizens' integrity – is incompatible with a liberal conception of justice. Then, I will reconnect the discussion of accommodations to my previous reflections and demonstrate that a plausible justification of some accommodation practices can rely on equal respect and toleration. I will further show that this shift from integrity to toleration and equal respect as the values justifying the legal recourse to accommodations is not of merely spec-

ulative interest, but leads, by contrast, to important normative implications regarding the types of complaints that can be legitimately be granted an accommodation for conscientious reasons.

To introduce the question of accommodations, consider the case of an always-diligent and obedient civil servant now faced with legislation that imposes a new duty on her that is conflicting with what she perceives as a compelling moral obligation of conscience.¹ She then asks to be exempted from the legal obligation. The society I see as embodying liberal toleration and equal respect is one where the problem of accommodating individual requests to be relieved from specific legal requirements is not solved by assessing the depth of the commitment the claimant is embracing. It is not the civil servant's integrity, I will argue, that fundamentally matters, on a liberal perspective, in this scenario, but rather her ability to pursue the conception of the good that she has autonomously chosen.

A citizen's ability to pursue her conception of the good is *unfairly* limited when the state enforces policies that have a different effect on different citizens. In this sense, any policy that diminishes some citizens' opportunity to pursue their conception of the good whilst allowing others to pursue their conception of the good without impediments *in a similar context* is, by itself, unfair. Alas, some cases of unfairness are unavoidable when it comes to distribution of opportunities. I will then argue that, insofar as the public authority can offer a valid justification, mere unfairness does not constitute a decisive reason to grant accommodations. I will point out, then, that a fairness-based argument in favour of accommodations can only succeed in case two conditions are jointly realised: the state cannot offer a valid justification

¹ The example is inspired by the story of Kim Davis, county clerk for Rowan County, Kentucky, who, after the decision in *Obergefell v. Hodges*, *US Supreme Court*, 2015, refused to comply with her legal job requirements and issue marriage licences to same-sex couples (<http://edition.cnn.com/2015/09/08/politics/kim-davis-same-sex-marriage-kentucky/index.html>).

for the norm that creates the unfairness, *and* the norm itself cannot be simply changed to prevent the unfair outcomes. I will show examples where both conditions obtain.

A more interesting case, however, is one where the state is in principle able to offer a justification for an unfair distribution of opportunities to pursue a conception of the good, which implies that no injustice would be committed in case the accommodation is not in fact granted. Still, I will argue that there are specific cases where reasons of toleration, of the type I defended in the previous chapter, would recommend granting the accommodation as a way of sustaining the long-term project of bringing about a stable liberal society. The accommodation would have value insofar as it fits into a pattern of virtuous tolerant policies whose central aim is to help build an environment where all citizens feel equally entitled to have their voices publicly heard and taken into account. Insofar as such toleration-based accommodations do not jeopardise the ultimate objective that the law is trying to achieve, and create only minimal burdens on third parties, they are not, I will argue, incompatible with a liberal theory of justice.

This is how the chapter is organised. In Section 1, I start by presenting the argument from integrity that we can retrieve in contemporary scholarship; I then argue that we can attribute a value to integrity either *conditionally* or *unconditionally*. The unconditional valuation of integrity faces two important objections: the objection from ‘villain’s integrity’, and the objection from the ‘elusive value of fixedness’. Whereas the first can be easily accommodated (for example by arguing that only ‘reasonable’ commitments matter for personal integrity), the second is more insidious. As lives devoid of integrity do not prevent people from developing crucial liberal virtues, I conclude that an unconditional valuation of integrity is incompatible with a broadly liberal theory of justice.

In Section 2, I show how some authors attempt to avoid this problem by reducing the value of integrity to a matter of merely subjective relevance; they render integrity valuable only for those who possess ‘deep commitments’ that they are not ready to sacrifice. The value they attribute to integrity becomes then conditional on the possession of ‘deep commitments’. That version of the argument, I argue, is flawed, because it conflates individuals’ subjective interests and their rightful entitlements. In Section 3, I show how a plausible response to the problem of how to treat (and sometimes accommodate) conscientious commitments that are incompatible with public norms can be organised following the principles of toleration and equal respect.

6.1 Liberalism and the Hopeless Valuation of Integrity

I will start with a definition of what I mean by integrity.² I will define integrity as the possibility to act in accordance with one’s *deep commitments*. This requires some further explanation. Firstly, integrity in this sense is inherently connected to actions, rather than conscience alone;³ integrity does not consist in the mere internal coherence of one’s belief set, but in the consistency between *some* beliefs and actions. Secondly, the beliefs of interest for integrity must be specially qualified.

The ‘deep’ beliefs that guarantee respect for integrity must be of particular significance to one’s conscience. This aspect, which will be particularly important

² The definition can be extrapolated from (Calhoun, 2016) (Bou-Habib, 2006, p. 117); (Ceva, 2010, p. 16); (Vallier, 2012, p. 155); (Laborde, 2015, p. 589); (Lenta, 2016, p. 247); (Vallier, 2016 b, p. 12) (Laborde, 2017 Chap. 6).

³ Notice that I am using ‘conscience’ in a purely descriptive sense, denoting each person’s deepest moral convictions. An alternative use of ‘conscience’ would be that of moral conscience as the awareness and understanding, on the part of each individual, of the requirements of morality. Kimberley Brownlee distinguishes the two usages (Brownlee, 2012, p. 2).

for this discussion, is the 'subjective relevance' of the beliefs that are of import to someone's integrity. How can we then characterise this 'subjective relevance'?

A possible answer is offered by Bernard Williams, in his observation that moral agents seem to place a special weight on some of their commitments, and would strongly object to seeing them balanced against the advancement of a presumed common interest. Even considering the balance would in fact 'alienate them from their actions and the source of their actions in [their] own convictions' (Williams 1980, p. 116). The agent might recognise that a certain course of action on her part will advance some common interest, but further believes that that course of action is beyond her own possibilities, not of course physically, but in the sense to do with the limits posed by her moral conscience.

To respond to Williams's challenge, the investigation will now take an excursion from political to genuinely moral theory, and in particular to that subdivision of moral theory which is the theory of values. We are trying to establish, to reiterate, whether, in a liberal regime, people are entitled to certain forms of legal accommodations on the grounds that strict compliance with the law might compromise their integrity. If we answer positively to this question, integrity is turned into a kind of liberal entitlement, on a par with, for example, freedom of speech or freedom of movement. Liberal entitlements presuppose a – however minimal – conception of the person; without at least some reflection on what it is like to have a flourishing (human) life, or what fundamental interests people have, we do not seem to possess the most basic conceptual resources necessary to answer the normative question about which basic freedoms or resources people are entitled to possess.

Hence, if we believe that integrity does constitute a liberal entitlement, we have to accompany that thought with the prior recognition that liberals ought to

value the preservation of integrity in somebody's life.⁴ I believe this recognition can come about in two different ways.

We place an *unconditional* value on integrity if we believe that a worthy and flourishing life *must* contain deep commitments that one is generally capable of respecting. We place a *conditional* value on integrity if we believe that, *for those who possess deep commitments* (a state of affairs we refrain from defining valuable in itself), the respect of those commitments constitutes a value.

Following Christine Korsgaard's analysis of the good (Korsgaard, 1983), we can then further distinguish, among the unconditional valuations of integrity, between an *intrinsic* and an *extrinsic* valuation. In the first case, we would value integrity for being in itself a source of value, in the second, for the value that it receives from another source, for example, human agents' positive attitude towards it. But, as Korsgaard herself vehemently argued,⁵ the belief in the existence of intrinsic goods, namely things that possess the property of goodness regardless of the positive dispositions that people might feel towards them, necessitates a cumbersome metaphysical apparatus. It equates, in fact, to a kind of metaphysical realism about moral properties, a position that, despite its significant popularity in the discussion of goodness at the beginning of the last century,⁶ few would be happy to take on board today.

⁴ Paul Bou-Habib follows a similar procedure, although with opposite results from mine, in his defence of accommodations in (Bou-Habib, 2006). Indeed, as he is mainly interested in accommodations for religious reasons, he first questions whether religion is itself a 'basic good' (pp. 111-114), before arguing that 'integrity' is instead the 'basic good' (p. 119) and that, because of that, 'the ability to perform [religious] conduct is derivatively valuable' (p. 110).

⁵ Both in (Korsgaard, 1983) and in the later (Korsgaard, 2013).

⁶ (Moore, 1922) (Ross, 2002 [1930]).

For an 'unconditional' valuation of integrity, therefore, I do not mean a valuation that takes the goodness of integrity to be *metaphysically* independent of human agents' positive disposition towards integrity. I mean, instead, a valuation that is not *neutral* regarding the inclusion of that positive disposition within one's value set. Valuing integrity unconditionally amounts to believing that the valuation of integrity is a rational choice that human agents cannot refrain from making, for integrity is, to reiterate, a necessary component in every human's flourishing and worthwhile life.

Each of the two ways of valuing integrity introduces distinct problems. An unconditional valuation must first include an explanation as to why a life containing deep commitments that are generally respected is valuable. A conditional valuation avoids this difficulty, but it meets the further and equally pressing complexity of having to fill the gap existing between the valueless status of possessing deep commitments and the valuable status of being capable of respecting, on an ordinary basis, those commitments.

Before I come to the analysis of the conditional valuation (which is prevalent in the literature on accommodations), the rest of this section is devoted to demonstrating why the unconditional valuation of integrity is incompatible with a broadly liberal political view.

Let's call my first objection the 'villain's integrity'. Imagine morality requires a specific moral agent to do something that she would consider deeply immoral, but that the vast majority of moral agents would consider a trivial action, if not 'the right thing to do'. Let's suppose, for instance, an ideologically committed racist is required by morality to help or protect a person belonging to an ethnic group he considers inferior, but he is disinclined to do so. Obnoxious as it might seem, it is possible to

stipulate the belief in the negative value of 'inferior' lives is a deep commitment for the ideological racist.

The objection applies to all uses of the concept of integrity that do not discriminate qualitatively, on a moral viewpoint, between the different commitments that compose a person's integrity. It is particularly insidious against any account that may justify legal accommodations by exclusive reference to respect for citizens' integrity, as the villain's integrity could never constitute a constraint on the implementation of justice, if this is to preserve internal consistency. Otherwise, one would reach the awkward conclusion that it is possible to identify the requirements of justice, but, insofar as there are still moral villains in the world, accommodating their views acts as a legitimate constraint on attaining justice itself.

The second objection tackles the unconditional value of integrity more directly. Williams's notion of integrity implies some commitments (what I have called 'deep commitments') are in some sort of way definitive of one's identity. This seems the most plausible way to interpret Williams's claim that '[i]t is absurd to demand of such a man [...] that he should just step aside from his own project and decision and acknowledge the decision [for the common good]' (Williams 1980, 116).

But, we may wonder, where does that 'absurdity' come from? Why should the mere fact of not seeing one's deep commitments thwarted be important enough to constitute an unconditional value? One might follow Williams again when he writes that certain commitments 'provide the motive force which propels [people] into the future, and gives [them] a reason for living' (Williams, 1981 b, p. 13). As so often with Williams, this sentence, despite its almost commonplace tone, contains a number of complexities and requires some interpretive exercise.

If the remark were meant to imply that people forced to compromise on their 'deep commitments' might as well terminate their life, we would have an extremely

strong reason to take integrity into account. But that seems implausible, as people who are forced into a choice between their allegiance to certain values and their compliance with certain norms still retain a plurality of options. One option might be a kind of retreat to the inner self, in search for that consistency between actions and conscience that the public norm is impeding. But that option might be unavailable in case the public norms, as in the example at the outset, is actively coercing someone into *acting* against his conscientious requirements. Still, one further option presents itself; slightly changing one's mind on the matter, and recognising that the conscientious requirement is one that it is possible, under certain circumstances, to disobey.

Then, a more plausible presentation of Williams's concern might be that certain commitments are so central to a person's identity that, if one were forced to part from them, he would transform into a different person, with a different personal identity. We can call this further characterisation of depth 'irreplaceability'; deep commitments are those an agent would not be capable of merely replacing, lest the entirety of his subjective value system inexorably collapses. This latter aspect might seem dramatic but, as Rawls famously argued, one of the very central features of a liberal theory of justice is that 'conversion is irrelevant to our public, or institutional, identity' (PL p. 32). If conversion is irrelevant to the way someone is publicly treated, there is no need to take active measures to avoid the scenario in which people, by giving up on their deep commitments, turn into a different type of believer and, due to the centrality of some of these beliefs, into a different type of person.

The value of the fixedness of one's 'deep commitments' then seems elusive: in any social context people are constantly required to treat at least some of their

commitments as flexible (or to see them as giving rise to merely *pro tanto* recommendations). Insofar as their public treatment is not conditioned by the commitments they entertain, this demand does not seem to produce excessive burdens.

Imagine now the life of a person who is happy to consider all her conscientious requirements as, to a certain extent, flexible. To give this figure more character, we suppose that she is a 'cosmopolitan' in the sense Jeremy Waldron used to describe Salman Rushdie. When confronted with the backlash after the publication of his *Satanic Verses*, Rushdie defended the work as celebrating 'hybridity, impurity, intermingling', rejoicing in 'mongrelisation' and fearing the 'absolutism of the Pure'.⁷ Now, as Waldron points out, the celebration of 'hybridity' and 'mongrelisation' is not the same thing as the embrace of autonomy, at least in the prevalent connotation given to the expression in post-Kantian moral philosophy. The 'autonomous' person is the one capable of both choosing a particular conception of the good and remaining substantially committed to it, whereas the 'mongrel' celebrated by Rushdie refuses any such choice of an 'absolute' conception.

What is missing in the cosmopolitan's life is an element of fixedness. The cosmopolitan does not seem to possess deep commitments, insofar as she does not attribute a particular predominance to any of the beliefs she happens to entertain. We might even complicate the character a bit more and add a post-modern element of 'irony' into her behaviour; the ultimate reason why the cosmopolitan does not seem to commit to any fixed value is that she cannot help questioning the validity of any minimally comprehensive conception of the good. In fact, she might even doubt that conceptions of the good may be subject to proper validity assessments beyond their merely 'sentimental' value; she will rank conceptions of the good, for instance, according to their displaying cruelty or encouraging solidarity.

⁷Quoted in (Waldron, 1992, p. 751).

Now, to be clear, my aim is not defending the cosmopolitan lifestyle, nor supporting the view of those (few) liberal thinkers who assume that the ironic, cosmopolitan character is more attuned to liberal philosophy than the rational autonomous one.⁸ Instead, my aim is to consider whether, starting from a broadly liberal perspective, we can consistently affirm that the cosmopolitan life, devoid as it is of deep commitments, is missing something of value.

Many liberal theorists would think that the cosmopolitan life is indeed missing something. They could claim, for instance, that the cosmopolitan, in refusing to choose a comprehensive theory of the good, is not exercising in full the Rawlsian moral powers. It might indeed seem that the cosmopolitan is not exercising in full her 'rational autonomy', i.e. her capacity 'to form, to revise, and to pursue a conception of the good, and to deliberate in accordance with it' (PL, p. 72).

The arguments I presented especially in chapter 4 should guard against this conclusion. Indeed, we should distinguish between the adherence to a fully comprehensive theory of the good, one that, in Rawls's words, expresses an 'intelligible view of the world' (PL, p. 59), and the mere exercise of the moral power to form and revise a conception of the good. The cosmopolitan is missing the opportunity to assent to a comprehensive and structurally articulated conception of the good, but might well form, through her infinite readjustments and responses to the most diverse cultural stimuli, a personal sense (to be distinguished from a full-blown conception) of the good. Moreover, the lack of fixedness does not prevent the cosmopolitan from 'deliberating in accordance' with her personal sense of the good; she can deliberate in this sense, we can argue, every time she takes a decision that is informed by her

⁸The argument is explicit in (Rorty, 1989), whereas I doubt Waldron's intent in is any different from mine.

(however unsystematic and flexible) sense of the good. It follows that the cosmopolitan's way of organising her own subjective values is a legitimate expression of the capacity to form and revise a conception of the good that, I argued previously, we ought to attribute universally if we want to respect our fellows equally.

Liberalism contains alternative notions of autonomy. In order to assess further the cosmopolitan lifestyle from a liberal perspective, it is therefore worth engaging with the perfectionist model of liberalism identifiable in the works of Joseph Raz and Steven Wall.⁹ Raz and Wall¹⁰ criticise traditional liberal neutrality on the grounds that, in an environment where people are exposed to a plurality of diverse and possibly incompatible values, the capacity to navigate that environment autonomously, by selecting the options one favours, is a fundamental human interest that the state should protect and promote.

If we follow Wall, however, we must conclude that nothing about the cosmopolitan's attitude suggests she is not 'charting [her] own course through life, fashioning [her] character by self-consciously choosing projects and taking up commitments from a wide range of eligible alternatives, and making something out of [her] li[fe] according to [her] own understanding of what is valuable and worth doing' (Wall, 1998, p. 128).¹¹ The cosmopolitan does not surrender her judgement to a pre-

⁹ Martha Nussbaum identifies the beginning of the liberal perfectionist project in Isaiah Berlin's rejection of the value-monism that characterised great part of the European moral philosophy, in favour of a view according to which 'there exists a plurality of overall accounts of how one should live, all of which are valid or objectively correct' (Nussbaum, 2011, p. 11). Raz indeed admits that his acceptance of Berlinian value-pluralism underpins his own defence of autonomy as the core liberal value (Raz, 1986, p. 133).

¹⁰ In (Raz, 1986) and (Wall, 1998) respectively. On a slightly different version of perfectionism, which consists instead in the identification of some natural goods that the state ought to legitimately promote, and which therefore is not grounded in complete value pluralism, see instead (Hurka, 1993) (Sher, 1997) and (Chan, 2000). It is to this latter form of perfectionism, and not to the former, that Rawls objects in *TOJ*, p. 325. See Richard Arneson's illuminating reconstruction (Arneson, 2000).

¹¹ See also (Raz, 1986, pp. 369-399). Raz uses the language of 'authorship' to qualify his account of personal autonomy; the autonomous person is the one who has both the means to exercise autonomy

determined, comprehensive conception of the good towards which she has no control, which would qualify her behaviour as heteronomous. Nor does she act irrationally by merely giving in to any whims and impulses that might cross her mind. On the contrary, she seems to conduct her life following values that she herself has deliberately chosen and partly readjusted – the only difference being that she does not seem to consider any of these values as permanently fixed and able to compose a comprehensive and structured conception. It might even appear that, in her constant readjustment of values in response to inputs coming from society, the cosmopolitan exercises autonomy more frequently than the average human being.

All these reflections are there to emphasise that there is a fundamental distinction between, on the one hand, integrity understood as the possession of and ‘deliberation in accordance’ with commitments one is *never* ready to sacrifice and, on the other, the exercise of autonomy that is usually taken as the model of human agency in liberal philosophy. The exercise of autonomy requires the ability to make significant, higher-order choices regarding the values that ought to inform one’s ordinary decisions in life. Being autonomous in this sense amounts to a capacity, in Harry Frankfurt’s terminology, to acquire ‘wholeheartedness’ (Frankfurt, 1988), to identify with the desires one entertains and to claim responsibility for their inclusion within one’s conscience. Nothing within this presentation implies – and this is the bulk of the distinction – that the autonomous agent must possess a particular hierarchy of values, with some of them (the ‘deep commitments’) so ingrained in one’s conscience that their replacement might compromise their holders’ identity.

Only the latter view, the unconditional valuation of integrity understood as the possibility of respecting one’s ‘deep commitments’, can sustain the normative

(‘mental abilities, an adequate range of options, and independence’ – p. 372) and does exercise that capacity with a sufficient degree of success.

claim that people are entitled to accommodations whenever such 'deep commitments' are frustrated. By contrast, a quasi-Kantian appeal to the values of autonomy or 'self-legislation'¹² will not serve the same purpose, for an agent's autonomy, understood as her capacity to set herself ends she might identify with, is not directly compromised when she is required by the law to take a course of action that partly conflicts with said ends. What the liberal respect for autonomy undoubtedly requires is that the public authority provides a justification for any legislative act that might affect citizens' chances to remain faithful to their chosen values.¹³ But this requirement of justification does not in any way preclude the possibility that the state might bring about a normative condition where some citizens find it more difficult than others to respect their (however 'deep' or 'shallow') commitments; it merely poses some constraints on its occurrence.

The upshot I derive from the above discussion is drastic; the unconditional valuation of integrity cannot be part of a liberal theory of justice. Firstly, unconditionally valuing integrity implies being incapable of discriminating between the commitments that are at least reasonable and those that are either morally obnoxious or harmful to others (we can categorise the racist's preferences under both rubrics). Secondly, the presentation of integrity as a liberal entitlement cannot be defended by treating integrity itself and autonomy interchangeably, for the two are clearly distinguishable concepts.

6.2 The Conditional Value of Integrity and Its Failure

The conclusion from the previous section was that the unconditional valuation of integrity is incompatible with a global commitment to liberal thinking. In this and

¹² As in (Ceva, 2010, p. 16).

¹³ I will elaborate more on this point in the next section.

the next section, I will determine whether a conditional valuation of integrity is more plausible and can therefore underlie a good argument in favour of accommodations.

As mentioned before, the main conceptual difficulty a conditional valuation of integrity encounters is that of filling the gap existing between the neutral status of possessing deep commitments and the valuable status of seeing those commitments generally satisfied. Theorists who adopt this view have to explain why the fact of acting consistently with one's own deep commitments, or the preservation of those commitments, has some additional valuable connotations that the mere possession of deep commitments alone lacks.

Two solutions to this puzzle dominate the literature. On the one hand, some authors reduce the value of integrity to its subjective relevance for those that possess deep commitments. This is the presentation of the argument I will address in the present section. Others avoid the reductionist move and show how integrity, far from being reducible to a single normative aspect, encompasses a plurality of moral concerns each of which might support a good argument in favour of accommodations. I will devote my last section to this presentation.

I will start this section by showing why subjective relevance cannot sustain by itself the 'argument from integrity'. The original argument holds, to reiterate, that the presence of an alleged threat on integrity can constitute a *pro tanto* reason to grant an accommodation. Hence, the argument from integrity in its 'subjective relevance' version runs like this: the subjective relevance, for some individuals, of their deep commitments, implies that the public authority owes them an accommodation whenever it requires them to act against said deep commitments. In this general formulation, the argument is flawed. It implies that the subjective relevance, for individual *y*, of whatever *x*, gives *y* a form of entitlement over *x*, such that, in case a

public authority requires *y* to hand over a part of *x*, *y* should be at least compensated. But that cannot be right; if I have an entitlement over *x*, that will mean I have a *right* to the possession of *x*, and this cannot be merely determined via the subjective relevance I place on *x*! I may well place a high relevance on things I have no right to possess (as a thief who grows fond of items she has stolen) and be indifferent towards things I have a right to (as a person who never goes under trial might feel towards the right to due process).

The argument therefore needs some further support in order to show that it does not lead to the previous, fallacious, results. That is why some authors tend to avoid the ‘argument from subjective relevance’ in its crudest form, and instead reduce the value of the integrity to a different moral category.

We may first see integrity as a significant component in people’s wellbeing, the one flourishing when people act in accordance with their perceived duties. This argumentative route is explicit in Paul Bou-Habib defence of a right to accommodations,¹⁴ whereas it implicitly underlies all those defences of accommodations that refer to the ‘centrality’, ‘speciality’ or ‘saliency’ of certain practices (mainly religious ones) or commitments to the people who pursue them.¹⁵ Even though the interpretation of integrity as a component of wellbeing is not always explicit, it is hard to imagine in which other regard these practices and commitments are ‘central’ to people’s lives.

¹⁴ ‘In order for all persons to enjoy an equal opportunity for well-being, I submit, each must be able to enjoy the basic good of integrity’ (Bou-Habib, 2006, p. 119).

¹⁵ See for example (Quong, 2006, p. 61); (Maclure & Taylor, 2011, p. 73); (Laborde, 2015, p. 597 ff.). In chapter 6 of her (Laborde, 2017), Cécile Laborde relies intensely on the notion of ‘saliency’, arguing in particular that the recognition of the ‘ethical saliency’ of certain interests (religious or otherwise) for specific communities of value can render them ‘legally salient’ (p. 202) within a liberal regime. I will have more to say about Laborde’s account in the next section.

The recognition of integrity as a component of wellbeing might plausibly give us *pro tanto* reasons to take into account people's deep commitments in practical deliberation. Once adjusted to address 'villain's integrity', these reasons could result in a normative requirement of the following kind; political authorities ought to accommodate integrity-based requests of non-compliance, provided this does not create comparable burdens on third parties' wellbeing.

As the value of integrity is here reduced to its relevance to people's wellbeing, it would indeed make sense to consider restrictions on third parties' wellbeing as the primary constraint on granting integrity-based accommodations. Granting an accommodation for reasons of wellbeing that resulted in restrictions on others' wellbeing would indeed confer on the latter class of citizens an unfair disadvantage.

However, questions about unfair advantages do not stop here. Even in case there is no apparent disadvantage for others, accommodating integrity-based requests for legal exemptions would still confer an advantage on those who can produce a specific kind of request (those who can say that their deep commitments are at stake). My take is that employing the language of wellbeing to justify such an advantage does not seem particularly promising.

I will explain this through an example. I take it that in at least some cases agents seek exemptions for conscientious reasons from laws that others would have an interest in violating, and that these different interests could all be potentially framed in terms of wellbeing. Take the infamous *Employment Division v. Smith* (494 U.S. 872 (1990)) case (American Supreme Court). The Court deliberated that Alfred Leo Smith, a member of the Native American Church, could not be exempted for religious reasons from prohibitions against drug consumption, despite the fact that peyote, the drug he was using, plays an important role in Native American religious

ceremonies. If the judges had employed integrity-based criteria, they would certainly have deliberated differently, but then their decision would have immediately triggered a fairness-based question; why is Smith exempted from the law, and not someone who simply enjoys peyote, but cannot cite any integrity-based reason to ground an exemption request?

What the court would have to say to those protesting the decision is simply that they, unlike Smith, are under no deep-commitment-based obligation to use peyote, and therefore their integrity was not compromised. This type of response should already sound suspicious to someone who has accepted my previous argument about equal respect as good-faith attribution of an equal capacity to form and revise a conception of the good. But I do not think one would have to buy my (controversial) account of equal respect to see that a defence of the value of integrity based on well-being is untenable. The peyote-amateurs excluded from the exemption could simply reply that integrity is, in the end, just *one* component in people's wellbeing, with the enjoyment of peyote for purely hedonistic reasons being another, and the reason the state is prioritising the first and giving no consideration to the latter is obscure.

I doubt the court could counter-reply that respect for one's perceived obligations constitutes an inherently *weightier* component in one's wellbeing. Why would it be necessarily the case that integrity is always the weightiest component? On a purely hedonistic and subjective conception of wellbeing, this is hard to show, and probably false. If one wanted to adopt, on the other hand, an objective conception of wellbeing, the reply would merely assume, without any argument, that integrity constitutes the most relevant component in wellbeing, regardless of the individual beliefs. We would be back, through such response, to the unconditional valuation of integrity that I maintain is incompatible with a liberal theory of justice.

Alternatively, we might reconnect the role of integrity to the classic liberal concern of public justification, from which I have started the present work. The argument can be summarised thus. A fundamental concern for all liberals is that, in order for an authority to be entitled to coerce its subjects into compliance, it has to be justified in terms that are acceptable to those citizens. Hence, if we simply extend the scope of public justification from the legitimation of the public authority to that of ordinary legislation, a move that, we have seen, is accepted by both ‘justificatory’ and post-Rawlsian political liberals,¹⁶ coercive norms that are justified in terms unacceptable to the coerced are illegitimate.

But what would make any particular norm’s terms of justification acceptable? I have already presented an answer that plays a prominent role in the ‘justificatory liberalism’ literature’: a law can legitimately require compliance only when citizens have ‘sufficient intelligible reason to endorse it’ (Vallier, 2016 b, p. 9). In the opposite case, when a law has been enacted following a legitimate procedure but is still unjustified to a number of citizens, who can express their complaint through intelligible reasons, an accommodation can then become the best option. Citizens who want to contest the justification of a particular law ought then to simply express their complaint through reasons that others can see have some relevance *for them*. As integrity is, Vallier assumes, a fundamental human concern, claims framed in terms of limits to one’s integrity are intelligible to the community at large.

Now, I have strongly criticised Vallier’s ‘intelligibility’ proviso in a previous chapter (§2.2), but, as I did before in my rejection of the reduction of integrity to well-being, I will not ground my critique of this type of reasoning in my previous arguments. I will assume, instead, that Vallier is right, to the extent that even the

¹⁶ See (Quong, 2011, pp. 273-287) and (Gaus, 2011, pp. 490-497).

cosmopolitan, who does not personally value integrity, can (intuitively?) see the relevance of integrity for others. What is still left unexplained, though, is the reason why, among all the types of reactive attitudes *vis-à-vis* someone else's possible loss of integrity, we should prioritise legal accommodations. Why couldn't we just feel sympathy, for instance, towards people that we can see will lose something relevant to them after the enactment of a law that compromises their integrity? This is, after all, a perfectly acceptable moral response in other, comparable cases. Take the convicted criminal's loss of freedom; we do not believe that in his case, considering the assumed justness of the sentence, we owe him more than sympathy, even though his loss of freedom is even more universally intelligible than a loss of integrity.

As such, 'intelligibility' is not to be conflated with 'acceptability'. I may well 'understand' my fellow citizens' loss of integrity as something relevant to them, and that even deserves some demonstration of sympathy, and still see it as a necessary price they have to pay to live in a democratic society where not everyone's preferences can be simultaneously satisfied. This point can be reformulated building on both of my previously presented objections. I may indeed understand the relevance *for you* of committing to a view I find repugnant, and even appreciate your commitment, and still refuse your request for accommodation. At the same time, I may still see how that value is *for you* fixed and non-negotiable, and still demand that you change your mind and accept a compromise, for compromising one's values is, in the end, an important virtue of co-existence.

A common limitation of these reductionist accounts is therefore their failure to provide an argument capable of moving beyond the focus on the subjective relevance of integrity. If, on the one hand, the subjective relevance of certain deeply held commitments is both trivial and undeniable, its role in a theory of justice is elusive.

Hence, accounts that simply derive implications on the right to legal accommodations from the subjective relevance of integrity are incapable of providing a response that is both reasonable and minimally persuasive to people who have an interest (but not one grounded in integrity) to violate a law for which others receive an accommodation.

6.3 The Elusive Role of Integrity

The previous section took issue with a particular way to explicate the ‘argument from integrity’. This section will consider instead an alternative presentation of the argument, in which the normative role of integrity is not reduced to its subjective relevance. A particularly promising suggestion is Cécile Laborde’s ‘Disaggregation Strategy’ in her analysis of religion. Instead of adopting a reductionist approach to integrity, Laborde shows how a multitude of different moral concerns can underlie our belief in the moral relevance of integrity.

Laborde’s ‘Disaggregation Strategy’ identifies two different ways through which integrity can have relevance for the individual. The first relevant sense of integrity she identifies is as a source of perceived moral obligations. People act with integrity in this first respect when they do not disobey those moral commands they have freely and conscientiously adhered to. The second is the sense more familiar to Williams; integrity is here taken to entail one’s identity in time due to the permanence of one’s commitment to the values one finds central to his or her conscience (Laborde, 2017, p. 215).

Starting from the belief in the centrality of both senses of integrity to individual agents, Laborde concludes that legal accommodations are not only acceptable but also demanded by justice, in two circumstances. The first circumstance is ‘disproportionate burden’ (pp. 221-229), occurring when the law imposes a burden on

some agents' pursuit of their conception of the good that is both weighty and removable with limited costs. The second is 'majority bias', the scenario in which the burden disadvantages a minority whilst allowing the majority to pursue their conception of the good without impediments in a somewhat similar context (pp. 229-237).¹⁷

When we analyse Laborde's substantive conclusions, however, we see that they are only partially grounded in the moral concerns she herself recognises as the bases of integrity. A mere appeal to the importance of perceived obligations or the preservation in time of personal identity does not show why 'disproportionate burdens' or 'majority bias' are those particular phenomena that require accommodations. We still need a further explanation as to why, given the moral relevance of the bases of integrity, we believe that these two specific cases are those in which the frustration of perceived obligations or personal identity demands a particular normative response. The view I will defend in these final paragraphs is that, once we understand what is morally relevant about these cases, and why accommodation is the appropriate response, the entire notion of integrity becomes redundant.

The cases at stake are, to reiterate, those in which citizens feel that their compliance with a specific norm would undermine their ability to live a life informed by principles and values they have autonomously adopted. In the Rawlsian language I have so far adopted, we can call this ability an opportunity to pursue one's conception of the good. Notice that this formulation does not presuppose that the conception of the good one intends to pursue must contain fixed and non-negotiable values.

¹⁷ In addition, she points out how the accommodations can be granted only after the requests for non-compliance have been subject to a test ascertaining their 'thick sincerity' (centrality to the individual's conscience) and 'thin acceptability', the latter of which ensures that the 'villain's integrity' objection is bypassed (Laborde 2017, 205-214).

The lack of focus on deep commitments, however, does not entail a complete exclusion of value-judgements and qualitative distinctions concerning individual preferences. The principle of equal opportunity to pursue one's conception of the good still entails a possible distinction between those preferences and beliefs that can be part of one's conception of the good and those, such as addictive or impulsive desires, that result from factors beyond one's control.

An example will help clarify this distinction. As I showed in the previous section, the 'deep commitments' model entails a different treatment, in the public sphere, between conscientious and hedonistic reasons; the first can give rise to legitimate accommodation requests, whereas the latter are excluded.¹⁸ On the 'conceptions of the good' model, by contrast, hedonistic reasons can be treated on a par with allegedly conscientious ones, as long as they can be shown to be part of one's conception of the good. So, if on the 'deep commitment' model we treat the religious believer differently from the hedonist, on the 'conceptions of the good' model we treat hedonists and religious believers equally. What we keep firmly separate, however, are preferences and beliefs that can form part of somebody's conception of the

¹⁸ In an analysis that bears some similarities with mine, Richard Arneson critically observes, 'affirming a right to moral freedom elevates the claim to accommodation of the conscientious above the claims of others that might have more compelling claims to accommodations' (Arneson, 2010, p. 1024). I do not entirely agree with the statement, as it seems to assume that 'conscientious claims' are necessarily less compelling than other, 'bread and butter' (ibid.) types of complaints. As will be evident in the next section, I do not intend to completely dismiss the role of conscientious claims in public discourse.

Similarly, my analysis from here on will differ from that of authors such as Peter Jones (Jones, 1994) or Simon Căbulea May (May, 2017) who have objected to exemptions for conscientious reasons on the basis of the Rawlsian 'social division of responsibility' between the state and citizens. My analysis, by contrast, helps showing why accommodating conscientious requests is sometimes either required by reasons of fairness or at least recommended for reasons of toleration.

good (be them conscientious or hedonistic) and preferences that are the mere result of addictions or impulses.¹⁹

Now, the question is, under which conditions can a norm hinder agents in their legitimate pursuit of a conception of the good? The first case I will consider is one in which an opportunity to pursue one's conception of the good is denied to some citizens and allowed to others *in a similar context and for unjustifiable reasons*. For a *similar context*, I mean a similar context in which different agents pursue their respective conception of the good; as conceptions of the good have (we empirically observe) a roughly similar structure,²⁰ the contexts in which agents tend to pursue them are generally comparable. Those might be contexts of actions, such as praying or resting or celebrating a holiday, or more general and less action-based contexts, that we can describe under comprehensive rubrics such as 'being faithful to one's values'. For 'unjustifiable reasons', I mean reasons that cannot be part of a reasonable theory of justice.

Now, a classic example of a scenario in which some citizens are denied opportunities to pursue one's conception of the good in this first sense is legislation about the weekly resting day in Western states. More specifically, we can say that Christians benefit from the public recognition of Sunday as the resting day in two different contexts; they can avoid working on that day, thus remaining faithful to a (currently very relaxed) principle of their religion, and they can have free time from

¹⁹ Thus, the distinction that matters under the 'conceptions of the good' model is not that between the hedonist and the conscientious believer, but that between the hedonist's desire to use peyote occasionally (as in the example above) and the addict's longing for another dose.

²⁰ In Laborde's terminology, we are here assuming that conceptions of the good are, to a certain extent, 'isomorphic' (Laborde 2017, 234). Laborde explicitly rejects that assumption, but I do not believe it is strictly necessary as long as we do not commit the mistake of taking as 'comparable contexts' purely identical concrete obligations (e.g. 'resting on Sunday' vs. 'resting on Saturday').

work for praying.²¹ Some words on why both advantages are unfair. Once we have agreed that opportunities to pursue one's conceptions of the good constitute an equal entitlement, then some justificatory reasons that would be acceptable in other circumstances (such as, for instance, appeals to the 'will of the majority' or the weight of tradition) become unavailable. It follows that we cannot cite the latter reasons as a legitimate justification for a restriction of someone's opportunities to pursue a conception of the good, in the same sense that we could not cite the 'will of the majority' as a legitimate reason to limit, say, freedom of expression.

To determine what was wrong in a scenario of this kind, we simply referred to the principle of fairness; fairness, in particular, regarding people's opportunities to pursue their conception of the good.²² A further condition is however necessary to grant an accommodation, if we want to avoid Brian Barry's well-known dilemma: either the law is unjust, and because of that must be amended, or is just, and as such cannot admit exceptions.²³ The further condition is that the law cannot indeed be amended, most of the time for pragmatic reasons, such as the fact that one law in this regard is necessary and, whatever its wording, will always end up disadvantaging some citizens' opportunities to pursue their conceptions of the good.²⁴

²¹ The first context puts Christians in an unfairly advantageous position *vis-à-vis* those religious believers, such as Jews or Seventh-Day-Adventists, who recognise a different sabbatical day. The second context gives Christians an unfair advantage *vis-à-vis* those religious believers, such as Muslims, who have a different day of prayer. The already mentioned *Sherbert v. Verner* (1972) case of the US Supreme Court addressed the first case, although without making use of the fairness paradigm, whereas the second context has been the object of the case *Ahmad v. United Kingdom* (1982) 4 EHRR 126 decided (against the plaintiff) by the European Commission of Human Rights (predecessor of the current European Court of Human Rights).

²² For an argument of this form, see (Quong, 2006).

²³ Barry premises the presentation of the dilemma with the adverb 'usually' (Barry, 2001, p. 40), and he indeed accepts that there might be justifiable cases of accommodations from just laws, although only for pragmatic reasons (Barry 2001, 49-54).

²⁴ Weekly resting day regulations are probably a good example. Indeed, a norm determining a universal resting day within a certain territory seems unavoidable and at the same time the identification of any day apart from Sunday would disadvantage Christians.

A second case is one in which, by contrast, the unfairness in the opportunity to pursue one's conception of the good is, at least *prima facie*, justifiable through reasons that are not in contrast with a reasonable theory of justice. Take the widely debated case of obligations to wear a crash helmet in risky situations (such as when riding a motorcycle or on a construction site), which contrasts with Sikh men's conscientious obligations to always wear a turban (*dastar*) to cover their long hair in public. Liberals that do not lean on the libertarian side would find the norm perfectly justifiable. The unfairness in the distribution of opportunities to pursue one's conception of the good, which the legislation undoubtedly creates, is therefore justifiable.

Still, as I tried to show in the previous chapter, the liberal ideology includes another principle that seems particularly apt to this case, that of toleration. In the previous chapter, I defended a particular *moral* model of toleration, contrasting it in particular to the view that sees toleration as primarily a *modus vivendi* considerations of civil harmony. Now we can see how the difference plays out in the practical context at stake.

In case of a model of toleration purely based on *modus vivendi*, the justification of accommodations will depend on a plurality of different contingencies. These can be, for example, whether Sikhs are particularly vocal in their expressing their opposition to the law, or whether they do generally ride motorcycles or work in construction sites, and so on. The assessment of these contingencies would be independent of any evaluation regarding the depth of the commitments of those requiring accommodation, but in a way that, I fear, would lead to a classic case of throwing the baby (representing here a value-based response to accommodations) out with the bathwater.

Of course, contingencies do matter on a political perspective, but there must be an underlying value justifying why, in some circumstances, the tolerant response to pluralism is preferable to one purely based on justice and fairness. Stability by itself cannot help, because a pure reliance on stability will lead to the unwelcome consequence that communities that are particularly vocal, if not particularly violent, will end up being privileged in the assignment of accommodations, as they are the ones producing a more credible threat. A result that, on top of being bad in itself, will also trigger the most perverse incentives, as people will be encouraged to always look more of a threat to public stability than they would probably like, as that will pay off in terms of being granted an accommodation.

On a genuinely moral model of toleration, by contrast, justifying accommodation will depend on whether the concession would help realise a particular moral ideal, which I tried to flesh out in the previous chapter. Again, I should repeat that the starting point for my reflections on toleration was the model society described by Scanlon in which citizens see each other as 'equally entitled to be taken into account in defining what [...] society is and equally entitled to participate in determining what it will become in the future' (Scanlon, 2003, p. 190). I suggest that, once we adopt this model of toleration, the question is not any longer whether certain norms risk compromising the peaceful co-existence of citizens living under the same political institutions, but whether the norms can create asymmetries among citizens concerning their capacity and efficacy in having a say on political decisions. Following this approach, one can reach the conclusion that a purely difference-blind application of abstractly justified norms can sometimes lead some minorities to feel, understandably, that their capacity to influence political decisions is gravely impeded. In that case, the state is asked to decide whether the understandable (although not necessarily legitimate) frustration of some citizens is a price that is necessary to pay

in order to maintain the commitment to the abstract principle of universal application of justified norms ('just laws admit of no exception'). The principle of toleration would suggest that, at least in some circumstances, and more than ever when the exception does not jeopardise the objective the law intends to promote,²⁵ the frustration could constitute a valid argument in favour of accommodations. This not because widespread frustration can lead to unrest in the polity, but because citizens who feel this form of frustration end up not enjoying the benefits of full membership in their society.

I am not sure that that the principle of toleration can support any more than some *pro tanto* normative recommendations, as it remains the case that sometimes the principle of universal application of justified norms ought to be prioritised. Its use here is, however, crucial, because it demonstrates that even laws that can be in principle justified, and that do not therefore create unjustifiable discriminations,²⁶ can nonetheless be subjected to exemptions if the objective is that of bringing about a society of equals.

The reflections in the previous paragraph serve the central purpose of questioning the role of integrity within arguments, such as Laborde's, that do not fall into the 'subjective relevance' trap. My conclusion about the role of integrity can only be, at this point, sceptical. We notice indeed how the two cases isolated by Laborde as those in which the recourse to accommodations is morally required can be captured by the two principles of fairness (more specifically, fairness in the distribution of

²⁵ A principle of toleration underlies, in my view, the balancing view that the American Supreme Court adopted following its decision in *Sherbert v. Verner*, 374 U.S. 398 (1963). What were balanced in the so-called 'Sherbert test' were, on the one hand, the significance of the burden suffered by the coerced party, and, on the other, the compellingness of the state interest. An exception from the law that did not clearly impede the general objective that a law was supposed to achieve (the 'state interest') would have been conceded.

²⁶ I am here using the concept of 'discrimination' in a purely descriptive manner, to denote all regulation that differently affects some of the agents that are subjected to it.

opportunities to pursue one's conception of the good) and toleration. To wit, 'majority bias' is a straightforward case of unfairness in the distribution of opportunities to pursue one's conception of the good, whereas 'disproportionate burden' is a case where the principle of toleration could have been easily applied but it wasn't, because the public authority chose to commit instead to the principle of the universal application of the law.

I come now to the advantages of my proposal. A major difficulty the literature on accommodations has had to address concerns the problem of determining *what* exactly the state is supposed to protect when it grants an accommodation for religious or conscientious reasons. A response that purely relied on religion as an independent category (namely, the state ought to protect faith or religious practice) would introduce a discrimination against holders of secular conceptions of the good that liberals could hardly defend.²⁷ On the other hand, an appeal to conscience would be more ecumenical, but would not by itself solve any conceptual or pragmatic puzzle. In order to understand exactly what kinds of requirements for accommodations are 'conscientious' (as opposed to, say, trivial or insincere) one would need first a full-blown theory of conscience;²⁸ a fundamental philosophical undertaking without any doubt, but one that might prove too onerous and controversial for the present task.

My account suggests instead that an argument in favour for accommodations can originate, from within a liberal perspective, every time a law leads to an unfair

²⁷ Indeed, many recent arguments in liberal political theory tend to de-emphasise the special role of religion and to defend religious freedom or tolerance (two classic liberal tenets) under more comprehensive moral categories, such as conscience, integrity itself or, with Ronald Dworkin, 'ethical independence'. See in particular (Schwartzman, 2012) (Leiter, 2013) (Dworkin, 2013) (Clayton, 2017).

²⁸ Kimberley Brownlee (Brownlee, 2012) certainly offers an interesting attempt in that direction, but her defence of conscience is primarily focused on the normative side, on whether citizens have a moral obligation to obey the law even when this goes against their conscientious requirements, or whether, as she eventually argues, they have a moral right to disobey in that case.

distribution of opportunities to pursue one's conception of the good (and cannot however be simply amended), regardless of whether the justification offered to support the norm is, in principle, legitimate. Whether the argument is legitimate will simply depend on whether the principles of fairness or toleration have been used appropriately, and on no more than that.

Granting a request for legal accommodation would not therefore require a costly and potentially controversial enquiry regarding the applicant's conscientiousness or the depth of her conviction; the state would remain neutral regarding the degree of affection that citizens might feel towards their conceptions of the good. The only types of enquiry that the liberal state is supposed to undertake, in deciding whether an accommodation ought to be granted, are the ones that are necessary to determine whether the applicant can legitimately appeal to either fairness or toleration. They concern such things as whether comparable contexts have been treated differently or similarly in the law (the fairness principle), whether possible differences in the treatment of comparable contexts are justified (the 'justifiable reasons' proviso in the fairness principle), or finally whether the objective that the norm sets out to achieve would be jeopardised in case specific exemptions were introduced, and whether, more generally, the accommodation would help in the achievement of a less divided society (the toleration principle).

A further advantage of my account is that it avoids the chance/choice distinction that has often haunted the discussion about cultural entitlements.²⁹ By grounding the concession of accommodations in the recognition of freedom to pursue one's conception of the good as an equal entitlement, the liberal state assumes that conceptions of the good are a product of citizens' autonomy. This assumption is of

²⁹ See for example Jonathan Quong's reply to Will Kymlicka's defence of cultural entitlements in (Kymlicka, 1995) (Quong, 2006, pp. 54-55).

course a simplification that abstracts away from the reality of religious and cultural affiliations, as well as from the importance of upbringing within certain cultural contexts. However, it is a simplification based on strongly egalitarian presuppositions; to prevent the state from engaging in those type of investigations about citizens' conscientiousness and sincerity that are likely to introduce controversies and frustrations in the polity, we treat all citizens – following my previous account of equal respect – *as if* they were indeed the authors of their conceptions of the good.

One final objection is worth considering. One could argue that, the last couple of paragraphs aside, a good deal of my argumentation has been devoted to proving that an appeal to the value of integrity is, in many cases, redundant. Admittedly, this 'fifth wheel' objection does not constitute by itself a decisive reason to abandon the notion of integrity altogether; parsimony in the use of abstract principles is just one consideration among many. One could argue that integrity, despite not playing a fundamental argumentative role, can still be helpful instrumentally within the public discussion. All things considered, the critic could continue, appealing to personal integrity is surely more understandable to the wider public than an abstract appeal to citizens' freedom to pursue their freely adopted conception of the good.

This objection is purely pragmatic and as such does not demand a full-fledged philosophical argument to be dismissed. Instead, my 'Closing Remarks' will present an example to show why this apparently common-sense consideration does not stand up to attentive criticism.

6.4 Closing Remarks on the Question of Accommodations

Removing integrity and replacing it with the less committal pursuit of one's conception of the good has not necessarily led to significant changes in the normative response to the problem at hand (at least *vis-à-vis* Laborde's approach), but has freed

the discussion from an unnecessary complication. The complication concerned the difficulty, expressed through my two objections, in either praising or at least giving a positive value to integrity starting from a broadly liberal perspective.

As I showed before, the main drawback caused by the normative focus on integrity was the unfair advantage attributed to those citizens that can express their complaints against a certain law in terms of limits on their fidelity to allegedly deep commitments. To make my point more persuasive, and answer the 'pragmatic' objection presented at the end of last section, I will conclude by noticing how those who cannot advance integrity-based complaints need not necessarily have a trivial reason to protest their exclusion from accommodation (such as, in the example above, their simply liking a drug used in ceremonial procedures). They could have, on the contrary, noble and altruistic reasons to oppose the law. Take the case of someone who believes military conscription is wrong (not only for himself, but for everyone), but who at the same time does not want to self-identify as a full-blown pacifist;³⁰ he simply finds a certain moral argument against military conscription overall persuasive. According to all the accounts hitherto mentioned, insofar as the agent cannot cite any deep commitment sustaining his opposition to the law, he cannot be granted an exemption, despite the fact that the state already recognises that an exemption from that specific legal obligation is justifiable (hence, in principle everyone could be exempted).

This example shows more than any other, I believe, the possibly darkest side of integrity-based accommodations; not only do they confer advantages that can be hardly justified, but they tend to disadvantage a particular kind of person. The disadvantaged are those that, in their opposition to the law, can only produce (good or

³⁰ 'Secular' pacifists have been granted conscientious objection from the draft in the United States since US Supreme Court sentences *United States v. Seeger*, 380 U.S. 163 (1965), and *Welsh v. United States*, 398 U.S. 333 (1970).

bad) arguments, unaccompanied however by the statement that the law undermines anything central to their identity. But this subtle disadvantage creates perverse incentives; more and more people will be motivated to abandon their flexibility (at least publicly) and act in a way that suggests their integrity, a concept previously completely stranger to them, is really at stake in cases where in fact it is not. The endemic use of the notion of integrity, instead of helping a polity come to terms with the possibility of legal accommodations, can give rise to a vicious circle in which more and more integrity-based accommodation requests are brought to the fore. In the end, however, those requests will be so insincere and hypocritical that, ironically, the meaning itself of integrity has got lost. To conclude, not only do 'deep commitments' have no special priority for justice, but an appeal to them in the public arena can be self-defeating.

Closing Remarks

In this final section, I recapitulate the argument and offer a series of points regarding what I take to be the main contribution of the work. These points aim at clarifying the relationship between my proposal and political liberalism.

The clarification is made necessary by the difference in content between the first and second half of the work. Whereas Chapters 1 to 3 jointly composed a critique of the political liberal treatment of legitimacy, Chapters 4 to 6 somehow led to a progressive disappearance of the theme of legitimacy, which came to be replaced by equal respect and toleration. Admittedly, equal respect and toleration, the values around which I build my own proposal, are distinct principles from legitimacy. The problem of disagreement in a society can have responses both on the basis of an account of legitimacy and an account of respect and toleration. Of course, my critiques to the political liberal model of legitimacy in the first half would still stand and that would imply I now need to provide an original account of legitimacy, something that, however, I have refrained from doing.¹ It appears I therefore need an explanation as to why legitimacy dropped out of the picture, and whether I conceive a plausible liberal response to the problem of disagreement that might not contain any reference to legitimacy.

It may also appear that I cannot avoid adopting an account of legitimacy given that I have proposed certain public policy measures, especially in the last chapter. Yet, the very possibility that certain policies may be enacted presupposes the existence of the state, and my observations at the beginning of Chapter 1 also disclose why I do not take anarchy to be an option. Hence, one may counter, if there is a state

¹ My scant comments about a promising theory of legitimacy at the end of Chapter 3 certainly do not amount to a full-blown account of legitimacy.

that is there to respond to problems generated by the existence of value-disagreement, shouldn't we make sure that that state is *legitimately* exercising its authority and coercion? In other words, shouldn't a solution to the problem of legitimacy somehow precede the identification of the policy instruments through which the state can adequately address pluralism?

In what follows I respond to these criticisms.

7.1 Liberal Pluralism beyond Legitimacy

I identified in the Introduction two sets of reasons that explain why political theorists of a liberal persuasion might be interested in value-disagreement. The first set of reasons is deontological in character; we want to investigate which type of respect, if any, is owed to people specifically in their role of public promoters of a diversity of values and ideas of the good. We notice that the circumstances of contemporary democratic policy, and of public deliberation in particular, force people to disclose something that might be crucial to their whole identity – namely, values they have adopted and that give content and meaning to their lives. This, as I noticed, would not be the case in a completely homogeneous society or in a society where deliberation is inexistent. This need to disclose and defend one's values and ideas of the good generates a question of respect: how can we make sure that, when people expose that part of their identity that consists of the values they most deeply cherish, they are not treated as inferior? This might well happen when, for example, certain conceptions of the good are called out as deviant or toxic or intolerably beyond the pale. Public deliberation creates, in sum, its own form of *vulnerability*: by demanding from citizens that they expose their values – at least when those values matter to the decision that is to be taken – it introduces the possibility that people might be disrespected precisely because of what they present of themselves publicly.

The second set of reasons is teleological. Here, pluralism is seen not so much as a source of vulnerability as a source of threats to civic co-existence. Pluralistic states are characterised by an often confrontational political dialectics; so confrontational that groups that perceive that their interests are often frustrated might be tempted to undertake a number of actions that, cumulatively, threaten the preservation of the authority of the state.

What interests, on the other hand, motivate the political liberal response to value-disagreement? As I have shown in the first three chapters, one theme in particular seems to dominate the political liberal literature: political legitimacy. The fundamental question informing political liberalism is, in brief, how to legitimise the authority of the state without relying on a set of reasons that, because of their being part of particular conceptions of the good, some citizens (namely, those that do not subscribe to such conceptions) might reasonably reject. We have also seen how this reconstruction can only capture one version of political liberalism. For the supporters of the more radical ‘convergence’ model, the problem is not so much justifying the authority of the state in general, but justifying the coercive imposition of particular norms.

A major conclusion of the present work is that both versions of political liberalism fail in two distinct ways. They first fail in the objectives they respectively ascribe to themselves, insofar as they produce accounts of legitimacy that suffer from inescapable ‘internal’ problems – problems I collected in the first three chapters and partly in the fifth.² There is no point in recapitulating those problems here. They also fail, however, because they seem to assert that the main problem generated by value-disagreement is that it constitutes an obstacle in the achievement of

²See in particular my critique to the attempts to use either public reason or ‘convergence’ as grounds for a viable account of political stability in §5.1.

political legitimacy or in the possible justification of coercive norms. In so doing, they do not respond to any of the concerns I have just set out above. .

To show why, imagine the critiques I expressed in the first three chapters were not decisive, and there was a political liberal way to reconcile legitimacy and disagreement. We would then have a legitimate state whose legitimacy is widely recognised (possibly, unanimously recognised) by citizens who adhere to a plurality of different conceptions of the good. Citizens would keep disagreeing about practically everything (including about which specific laws are right and wrong), but would all converge on the belief that the state can legitimately demand that they adjust their conduct following its directives.

Is a scenario of this kind even feasible? It seems highly unlikely that citizens might accept the authority of the state not because they give a generally positive evaluation to its laws, but because they recognise that the state is – according to certain abstract criteria – legitimate.³ But suppose this is indeed the case. Still, citizens are not protected by at least the first of the problems I previously identified – they would still be vulnerable to that particular form of disrespect that involves being discriminated against or marginalised because of the subjective values they publicly advance. Insofar as the public culture is one that encourages the public expression of subjective values (something that is valuable all-things-considered), it is of small consolation to citizens to know that the justification of public authority is not in principle incompatible with their values, when the expression of the same values is often greeted with scorn or incredulity or, worse, requests for containment.

³I provided an alternative representation of what it might mean for citizens to see the authority of their state as legitimate in §2.3.2.

By saying it is 'of small consolation', however, I have conceded even too much. Because in fact the public justification of the authority is of no consolation whatsoever when the problem is that of demonstrating respect to citizens in their particular role as value-promoters. Is it possible to imagine a condition where citizens are routinely disrespected and yet the state is legitimate? The answer would depend, of course, on which model of legitimacy is adopted. If we take the model of legitimacy prevalent in the Rawlsian conception of political liberalism, where legitimacy is reduced to public justification to the (idealised) reasonable, it seems that nothing prevents the emergence of a scenario where a) the state is legitimate (because of the justification of its public authority meeting the prescribed criteria) and b) citizens do not receive – either from the state or from their fellows – the required level of respect.

One might counter this by arguing that 'public justification to the reasonable' is an ideal that must include some reference to the equal respectful treatment of the citizens. Admittedly, the liberal principle of legitimacy is only appealing to reasonable citizens who, by definition, subscribe to a certain view of equality, namely to a view of their fellows in society as free and equal participants to the 'system of social cooperation' that the state is taken to be. But this commitment to equality can be plausibly interpreted to simply entail the constitutional entrenchment of a more or less robust scheme of basic rights and liberties, the same Rawls defends as parts of each society's 'constitutional essentials' (PL, 227-230) or as grounding a state's sovereignty in the international arena (Rawls, 1999, p. 79). Whether the implementation of such rights would lead to the emergence of a public culture of equal respect, however, is still to be demonstrated.

By emphasising the distance between the political liberal account of legitimacy and the recognition of equal respect I am not in any way suggesting that a state

is legitimate only if it guarantees that all its citizens receive the form of equal respect I have introduced in Chapter 4. Again, this is one question only a full-blown account of legitimacy could answer. What I am suggesting, on the other hand, is that the question of the appropriate response that is due to the vulnerability generated by value-disagreement and public deliberation is completely separable from political legitimacy. Of these two questions, only the one about the vulnerability of value-disclosure is a direct output of value-disagreement (combined, it is worth pointing out, with democratic deliberation) whereas the question of legitimacy can at most have an answer that is particularly sensitive to the existence of disagreement (this, I take it, is the fundamental insight of political liberalism), but does not presuppose value-disagreement to be formulated. Indeed, the duty to legitimise its authority would apply to states regardless of whether their public culture is one of pluralism or of homogeneity. One might argue that the circumstances of pluralism somehow constrain the strategies that the state can appeal to in the justification of its authority, but this does not prove any sort of causal link between the emergence of pluralism on the one hand and the emergence of legitimacy as a condition that the state must satisfy on the other.

Similarly, I have suggested in Chapter 5 that the two questions of legitimacy and stability do not entail one another. Which, however, does not imply that stability is not a moral notion, as the most radical supporters of what I have called the 'grim view' behind political realism would probably suggest. On the contrary, I have tried to show that only a robust commitment within society to the virtue of toleration can solve the question of stability. But, if toleration and stability are then intricately related, at least in the contemporary liberal world – because only a society where the virtue of toleration prevails can aspire to stability – the same cannot be said of legitimacy and stability.

Does this imply that a legitimate society can be unstable or, alternatively, that a stable state can be illegitimate? Let's analyse both horns of the question, setting aside, as I did before, my critiques to the political liberal model of legitimacy. As I showed in Chapter 5, nothing about either the political liberal notion of public reason or the 'political justificatory' notion of convergence seems to guarantee that states whose authority is justified will remain stable. The demonstrations political liberals and supporters of 'justificatory liberalism' alike offer for the mutual entailment of legitimacy and stability is unpersuasive and self-defeating. Legitimacy and instability can therefore, under the political liberal model, co-exist.

But we can also notice how the problem of stability affects legitimate and illegitimate societies alike. This might seem a trite point. Rawlsian liberals could simply respond by conceding that illegitimate societies can be stable or unstable, but what they are interested in is 'stability for the right reason', that is, the stability that obtains when citizens recognise that their allegiance to the state does not in any way compromise the attachment to the subjective values they have autonomously adopted. But this semantic answer ('you take stability to be what ordinary language generally labels 'stability', we give it a technical meaning') is implausible.

Firstly, 'stability for the right reasons', as we have seen, is a much less appealing model than Rawls's initial presentation in *Political Liberalism* might indicate. Once we start enquiring about what it means to use public reason as a tool in the achievement of stability, we see that the signals citizens send to each other by embracing public reason hardly demonstrate that they intend to keep complying with public norms even when this goes against their interest.⁴ Until additional empirical evidence is brought to bear on this matter, the point is not fully supported. Hence,

⁴ The same can be said for the signals sent by citizens who engage in the 'convergence' procedures defended by 'justificatory' liberals.

as I wrote before for the political liberal notion of legitimacy, even the account of stability that prevails within political liberalism (of both versions) does not satisfy its own standards of success.

There is a further reason, however, to reject the ‘semantic’ answer about stability, which is that one ought to be cautious about completely discounting the positive role of the stability of illegitimate states. If we agree there is a certain value in stability – and not just in ‘stability for the right reasons’ – then that value might persist whether or not we face a legitimate state. The question is to see, therefore, whether there is a value in political stability that is different from the merely derivative value that the preservation of a *valuable* legitimate state might possess.

This question – about the possible value of the stability of *certain* illegitimate states – must be neatly distinguished from the different one about the value lying in the stability of the state in general, regardless of how it treats its citizens or exercises its authority. To the latter question, it seems we can only give a negative answer; there is no value in the preservation of a deeply unjust state. But there might value in the stability of states that do not meet certain criteria of legitimacy.

As political stability is defined as the preservation of both the authority of the state and the unity of the polity, the value of stability must be identified in the value of having an authoritative and unitary state. I have already excluded, in Chapter 4,⁵ one possible source of value for the state. In criticising the idea that self-determination and freedom can justify an entitlement to the preservation of certain communities of value, I also exclude the option that the freedom or self-determination of a people might justify the preservation of a state, irrespective of the way in which it exercises its authority.

⁵ See my rebuttal of Will Kymlicka’s model of liberal multiculturalism in §4.2.2. Similar arguments were also used to dismiss the liberal nationalism of David Miller in §1.

There is another sense, however, in which the state might be said to be valuable, and it derives from the Kantian argument about the necessity of the civil authority as a vehicle for the substantive protection and enforcement of those rights that, outside the authority of the state, are merely ‘formal’ and ‘indeterminate’ and depend for both their determination and implementation on the good will of uncoordinated moral agents. A state is needed, Kantian interpreters affirm,⁶ to fully guarantee respect for every citizen’s moral ‘independence’, for their capacity, that is, to determine and try to pursue their ends in a manner that is compatible with others in society doing the same.

In the (Rawlsian) parlance I have employed so far, the state is needed in order to guarantee that people can freely employ their moral capacity to form, revise and look after an autonomously chosen conception of the good without impinging on others doing the same. The liberal state works therefore in the present investigation as both a source of vulnerabilities and as a possible solution to such vulnerabilities. We have seen that public deliberation in particular is a source of vulnerability, insofar as it demands from citizens that they participate to certain debates where they might be forced to disclose some of their personal values and preferences. A culture of equal respect, I have argued, is then needed in order to prevent such disclosures from becoming a constant source of hostility. But the main entity that can encourage the emergence of such a culture is the state, primarily through the enforcement of a robust scheme of equal rights and liberties – most specifically, the most extended freedom of expression that is compatible with respect for the dignity and bodily integrity of all citizens.

Freedom of expression is not sufficient to guarantee that the culture of public deliberation turns into one of equal respect – citizens can still disrespect each other

⁶ See in particular (Waldron, 1996) (Ripstein, 2004) (Pallikkathayil, 2016).

whilst allowing others to express themselves freely. Nonetheless, it seems undeniable that one particularly pernicious way to exploit the vulnerability generated by value-disclosure is to pre-empt the disclosure of certain modes of expression or – even worse – to forbid such disclosure to a particular set of citizens. Regardless of how vulnerable people might feel at having to expose parts of their identity, taking active measures to limit such exposure to certain categories of people sends one of the most disrespectful signals to those at the receiving end – namely, that they are incapable of exercising their capacity to form and revise a conception of the good in a way that is worthy of respect, and that its exercise can therefore be legitimately impeded.

I wrote before that people can still disrespect each other under conditions of freedom of expression. Under the account of equal respect I defended in Chapter 4, that is most certainly the case. Indeed, freedom of expression is a pre-condition for the exercise of equal respect but does nothing by itself to guarantee that citizens who participate the public debate receive the form of respectful treatment I defended there. Which is why a state must then be involved in more than just enforcing freedom of expression if a culture of equal respect is to emerge.

Let me then recapitulate briefly what are the strings of policies I defended throughout the work, all of which have as their core aim and ambition the promotion of equal respect or toleration.

1. In §6.3, I argued that, if a state needs to assign recognition to particular communities of value – for example by substantially advantaging the exercise of the capacity to form and revise a conception of the good for their members – then citizens who do not belong to that community must be compensated through a process of accommodations. And

this not for generic reasons of neutrality but for reasons of equal respect, which impose that, if the state offers recognition to certain communities of value, it must do so in a way that does not lead to disrespectful implications for citizens that do not belong to that community.

2. In §4.2.2 I attacked the idea that particular communities are entitled to some form of assurance, from the part of the state, regarding their preservation and prosperity. Again, a state's commitment towards the preservation of certain communities would be consistent with equal respect only in the implausible circumstances where that the state is capable of guaranteeing prosperity to all the (present and nascent?) communities of value existing within its territory.
3. In §5.3.3-4, I suggested that the promotion of justice itself can play a crucial role in the emergence of a public culture of toleration. In particular, the more a state approximates a condition of justice, the less it will contain communities of citizens who, because of their interests being constantly frustrated, will understandably feel mistrust for the rest of the polity.
4. Finally, in §5.3.4 and §6.3, I remarked how, to further enhance the spread of a culture of toleration, majorities should be willing to strike a compromise even when this is not strictly required by prudential reasons of self-preservation. Accommodations can be granted to minorities not only for reasons of fairness but also in order to send two important signals to minorities. The first is that their conception the good is not one the state refuses to even consider worth pursuing or that people should be somehow ashamed of publicly displaying. The

second is that, although the group in question was in this occasion outvoted (and that fact alone is what makes it a minority), its role in contributing to the definition of the society remains unaffected.

These strings of policies constitute, together, the sketch of a liberal response to the question of value-disagreement. What makes them liberal is that they are organised around the core liberal values of equal respect and toleration. What makes them responsive to value-disagreement, in a way that I take to be superior to the focus on legitimacy and neutrality which characterises political liberalism, is that they directly respond (with more or less efficiency depending on the empirical contextual data I keep excluding from this analysis) to the two problems I have previously identified as constitutive of the 'question of pluralism'. A public culture of equal respect – promoted in particular through 1. and 2. – protects citizens from the vulnerability generated by value-disclosure. A public culture of toleration – which 3. and 4. aim at enhancing – increases the chance that a diverse society can remain stable, thus preventing that escalation from diversity to constant hostility that the, under the 'grim view', only the most severe Leviathan can avoid.

It remains to be shown, however, what is the role of political legitimacy in this picture. In Chapter 1, I defended a model of legitimacy as the right for a state to exercise authority, or, to put it differently, to demand that citizens adjust their conduct following the directives issued by the states. As the policies alluded to in the previous list are all authoritative (and coercive), it might then follow that they could not be implemented by illegitimate states, because the illegitimate state would have no right to demand that citizens adjust their conduct following them.

Hence, a possible objection to all the theses presented so far is that my argument is incomplete, insofar as I have not defended a *standard* (as opposed to a

simple account – see my distinction in §1.1) of legitimacy. This objection is correct – up to a point – but not decisive. It is partly correct because, in the absence of a standard of political legitimacy, I have not excluded the possibility that all states in the real world could fail to be legitimate, thus compromising the likelihood that the policies that aim at promoting equal respect or toleration might be implemented.

However, citizens might have reasons to comply with certain norms even when the state itself has no right to demand that they comply – for example, if the norms are ones of justice. Similarly, those that are contingently capable of exercising power in an illegitimate state, even if they do not have a right to command, are under a moral obligation to exercise that power in a way that is as close as possible to justice. So, if they are contingently capable of issuing directives that those subject to their power tend to obey, they should make sure that said directives promote justice. As I am confident that the policies mentioned before are ones that, at least under favourable conditions, would promote justice, sovereigns and subjects alike would be under an obligation to – respectively – implement and respect them even in conditions of political illegitimacy. Of course, citizens would not in that case owe respect of said norms to the state, because the state could not demand their obedience. But, in case the norms were implemented, citizens would not be free to ignore them, in the same way as, even under the cruellest autocracy, citizens would still be under an obligation to respect just norms.

Secondly, legitimacy is only one of the conditions of successful implementation of state policies. Other conditions I have not treated here are, for example, the existence of a functioning enforcement mechanism or the presence of a public culture of compliance – although this latter aspect is obviously one of the elements of stability. This to explain that my account is certainly incomplete, but in a way that I

have admitted from the beginning and that has to do, ultimately, with the necessary limits imposed by the methodology of political theory.

Finally, and most importantly, I am not persuaded by the main thesis of political liberalism – namely, that a standard of political legitimacy for contemporary democratic societies *must* be sensitive to the fact of value-disagreement. Imagine an account of legitimacy that holds that a state is legitimate if its institutions are organised in a way that promotes and maintains the justice of its ‘basic structure’.⁷ Imagine further that we define justice in a way that abstracts from or even ignores value-disagreement, for example arguing that justice is explained in terms of aggregate average utility or priority for the worse-off or equality in the distribution of primary goods or that it corresponds to a state of affairs where sheer luck has no effect on the chances of individuals, and other such accounts.

What is wrong with that definition of legitimacy? The canonical political liberal answer is that this way the authority of the state is justified in a way that is not appealing to every citizen who embraces a reasonable conception of the good. But this answer is in one way false and in another question-begging. It is false because nothing prevents people who embrace a plurality of reasonable conceptions of the good from accepting a particular account of justice as persuasive. Arguably, none of the previously mentioned conceptions of justice would be appealing to a theocratic religious believer, but I guess a theocratic believer would not in any case qualify as reasonable. But nothing in principle prevents, for instance, a non-theocratic Muslim and the proverbial nihilistic surfer from both accepting, say, luck-egalitarianism or prioritarianism as a plausible account of distributive justice. Hence, that an account of legitimacy that is not sufficiently sensitive to pluralism would lead to a justifica-

⁷I have alluded to this model at the end of Chapter 3.

tion of the authority of the state that is inaccessible to a significant number of reasonable citizens is still to be demonstrated. Secondly, as I argued extensively in Chapter 2, simply assuming that legitimacy must include an appeal to public justification runs the risk of turning public justification into some kind of political liberal fetish. It is not obvious what foundational moral value might underpin the political liberal interest in public justification.

If, however, the liberal political attempt at defending its own standard of legitimacy had at least succeeded, I would have to explain why I would decide to ignore it and to suggest instead that a political liberal response to value-disagreement can neglect the question of legitimacy. But I have dedicated three entire chapters to collecting all reasons explaining why this attempt ultimately fails. Other attempts at reconciling legitimacy and value-disagreement might of course be more successful, but, as I wrote before, the very point of a pluralism-sensitive account of legitimacy seems unmotivated, or at least insufficiently motivated. The necessary connection between sensitivity to pluralism and a plausible account of legitimacy still needs elucidation, given that 1) sensitivity to pluralism does not automatically lead to public acceptance of the authority of the state ('justification to the reasonable' is still an idealisation!) and, 2) *vice versa*, pluralism-insensitive accounts of legitimacy can still be accessible to and persuasive for a plurality of citizens that endorse the most diverse conceptions of the good.

Political liberalism sees the 'problem of pluralism' to lie in the difficulty of presenting an account of legitimacy that a plurality of citizens who endorse a host of different and reasonable conceptions of the good might accept. I reply that this project cannot succeed and is unnecessary. It cannot succeed, for the reasons I presented in the first three chapters, because it is incapable of capturing the phenomenon of value-disagreement in its nuances and complexities. But it is unnecessary, as

I tried to show in the last three chapters, because there are more pressing and relevant questions that are generated by pluralism, questions that an account of legitimacy would be ill-suited to respond to. As this thesis has argued, the liberal response to disagreement can only succeed by overcoming its restricted focus on legitimacy.

Bibliography

- Abizadeh, Arash, 2004. Historical Truth, National Myths and Liberal Democracy: On the Coherence of Liberal Nationalism. *The Journal of Political Philosophy*, 12(3), pp. 291-313.
- Anderson, Elizabeth., 1999. What is the Point of Equality?. *Ethics*, 109(2), pp. 287-337.
- _____, 2010. *The Imperative of Integration*. Princeton (NJ): Princeton University Press.
- Arneson, Richard, 2000. Perfectionism and Politics, *Ethics*, 111(1), pp. 37-63.
- _____, 2010. Against Freedom of Conscience. *San Diego Law Review*, 47(4), pp. 1015-1040.
- Aumann, Robert J. and Sergiu Hart, 2003. Long Cheap Talk. *Econometrica*, 71(6), pp. 1619-1660.
- Baderin, Alice, 2014. Two Forms of Realism in Political Theory. *European Journal of Political Theory*, 13(2), pp. 132-153.
- Balint, Peter, 2006. Respect Relationships in Diverse Societies. *Res Publica*, 12(1), pp. 35-57.
- _____, 2015. Identity Claims: Why Liberal Neutrality Is the Solution, Not the Problem. *Political Studies*, 63(2), pp. 495-509.
- Barry, Brian, 2001. *Culture and Equality*. Cambridge (Mass.): Harvard University Press.
- Beccaria, Cesare, 1995 [1764]. *On Crimes and Punishments*. Cambridge: Cambridge University Press.
- Bedi, Sonu, 2007. Debate: What is So Special about Religion? The Dilemma of the Religious Exemption. *The Journal of Political Philosophy*, 15(2), pp. 235-249.
- Beitz, Charles, 2001. Human Rights as a Common Concern. *The American Political Science Review*, 95(2), pp. 269-282.
- _____, 2009. *The Idea of Human Rights*. Oxford: Oxford University Press.
- Bellamy, Richard and Martin Hollis, 1995. Liberal Justice: Political and Metaphysical. *The Philosophical Quarterly*, 45(178), pp. 1-19.
- Benn, Stanley I. and W.L. Weinstein, 1971. Being Free to Act, and Being a Free Man. *Mind*, 80(318), pp. 194-211.
- Bentham, Jeremy, 1970 [1780]. *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press.

- Beyleveld, Deryk, 2014. Human Dignity and Human Rights in Alan Gewirth's Moral Philosophy. In: M. Düwell, J. Braarvig, R. Brownsword and D. Mieth, eds. *The Cambridge Handbook of Human Dignity*. Cambridge: Cambridge University Press, pp. 229-239.
- Billingham, Paul, 2017. Convergence Liberalism and the Problem of Disagreement Concerning Public Justification. *Canadian Journal of Philosophy*, 47(4), pp. 541-564.
- Bou-Habib, Paul, 2006. A Theory of Religious Accommodations. *Journal of Applied Philosophy*, 23(1), pp. 109-126.
- Brandt, Richard, 1964. The Concepts of Obligation and Duty. *Mind*, 73(291), pp. 374-391.
- Brennan, Geoffrey, 2001. Collective Coherence?. *International Review of Law & Economics*, 21(2), pp. 197-211.
- Broadie, Alexander and Elizabeth M. Pybus, 1975. Kant's Concept of 'Respect'. *Kant-Studien*, 66(1), pp. 58-64.
- Brownlee, Kimberley, 2012. *Conscience and Conviction: The Case for Civil Disobedience*. Oxford: Oxford University Press.
- Buchanan, Allen, 1999. Recognitional Legitimacy and the State System. *Philosophy & Public Affairs*, 28(1), pp. 46-79.
- _____, 2002. Political Legitimacy and Democracy. *Ethics*, 112(4), pp. 689-719.
- _____, 2011. Reciprocal Legitimation: Reframing the Problem of International Legitimacy. *Politics, Philosophy & Economics*, 10(1), pp. 5-19.
- Calhoun, Cheshire, 2016. Standing for Something. In: *Moral Aims: Essays on the Importance of Getting It Right and Practicing Morality with Others*. Oxford: Oxford University Press, pp. 123-154.
- Carter, Ian, 2011. Respect and the Basis of Equality. *Ethics*, 121(3), pp. 538-571.
- _____, 2013. Are Toleration and Respect Compatible ?. *Journal of Applied Philosophy*, 30(3), pp. 195-208.
- Ceva, Emanuela, 2010. Self-legislation, Respect and the Reconciliation of Minority Claims. *Journal of Applied Philosophy*, 28(1), pp. 14-28.
- Chan, Joseph, 2000. Legitimacy, Unanimity, and Perfectionism. *Philosophy & Public Affairs*, 29(1), pp. 5-42.
- Christensen, David, 2007. Epistemology of Disagreement: The Good News. *The Philosophical Review*, 116(2), pp. 187-217.

- _____, 2009. Disagreement as Evidence: The Epistemology of Controversy. *Philosophy Compass*, 4(5), pp. 756-767.
- Clayton, Matthew, 2017. Is Ethical Independence Enough?. In: C. Laborde and A. Bardon, eds. *Religion in Liberal Political Philosophy*. Oxford: Oxford University Press, pp. 132-146.
- Cohen, Andrew Jason, 2004. What Toleration Is. *Ethics*, 115(11), pp. 68-95.
- Cohen, Gerald Allen, 2003. Facts and Principles. *Philosophy & Public Affairs*, 31(3), pp. 211-245.
- _____, 2009. *Why Not Socialism?*. Princeton (NJ): Princeton University Press.
- _____, 2011. Rescuing Conservatism. A Defense of Existing Value. In: R. J. Wallace, R. Kumar & S. Freeman, eds. *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*. New York: Oxford University Press, pp. 203-226.
- Cohen, Joshua, 2006. Is There a Human Right to Democracy?. In: C. Sypnowich, ed. *The Egalitarian Conscience: Essays in Honour of G. A. Cohen*. Oxford: Oxford University Press, pp. 226-246.
- Copp, David, 2008. The Idea of a Legitimate State. *Philosophy & Public Affairs*, 28(1), pp. 3-45.
- D'Agostino, Fred, 1996. *Free Public Reason: Making It Up As We Go*. New York : Oxford University Press.
- Dancy, Jonathan, 1993. *Moral Reasons*. Oxford: Blackwell.
- Darwall, Stephen, 1977. Two Kinds of Respect. *Ethics*, 88(1), pp. 36-49.
- _____, 2004. Respect and the Second-Person Standpoint. *Proceedings and Addresses of the American Philosophical Association*, 78(2), pp. 43-59.
- Dworkin, Ronald, [1986] 1998. *Law's Empire*. Oxford: Hart Publishing.
- _____, 1989. Liberal Community. *California Law Review*, 77(3), pp. 479-504.
- _____, 1997 [1970]. Taking Rights Seriously. In: *Taking Rights Seriously*. New York: Bloomsbury, pp. 223-247.
- _____, 2011. *Justice for Hedgehogs*. Cambridge (Mass.): Belknap Press of Harvard University Press.
- _____, 2013. *Religion Without God*. Cambridge (Mass.): Harvard University Press.

- Edmundson, William E., 1998. *Three Anarchical Fallacies*. New York: Cambridge University Press.
- Elga, Adam, 2007. Reflection and Disagreement. *Noûs*, 41(3), pp. 478-502.
- Enoch, David, 2010. Not Just a Truthometer: Taking Oneself Seriously (but Not Too Seriously) in Cases of Peer Disagreement. *Mind*, 119(476), pp. 953-997.
- _____, 2013. The Disorder of Public Reason. *Ethics*, 124(1), pp. 141-176.
- _____, 2015. Against Public Reason. In: D. Sobel, P. Vallentyne and S. Wall, eds. *Oxford Studies in Political Philosophy. Volume 1*. Oxford: Oxford University Press, pp. 112-140.
- _____, 2017. Political Philosophy and Epistemology. The Case of Public Reason. In: D. Sobel, P. Vallentyne & S. Wall, eds. *Oxford Studies in Political Philosophy. Volume 3*. Oxford: Oxford University Press, pp. 132-165.
- Erman, Eva and Niklas Möller, 2015. Political Legitimacy in the Real Normative World: The Priority of Morality and the Autonomy of the Political. *British Journal of Political Science*, 45(1), pp. 215-233.
- Estlund, David, 2008. *Democratic Authority: A Philosophical Framework*. Princeton (NJ): Princeton University Press.
- Feinberg, Joel, 1970. The Nature and Value of Rights. *Journal of Value Inquiry*, 4(4), pp. 243-260.
- Feldman, Richard, 2009. Evidentialism, Higher-Order Evidence, and Disagreement. *Episteme*, 6(3), pp. 294-312.
- Flikschuh, Katrin, 2010. Kant's Sovereignty Dilemma: A Contemporary Analysis. *The Journal of Political Philosophy*, 18(4), pp. 469-493.
- Foot, Philippa, 2002. Goodness and Choice. In: *Virtues and Vices. And Other Essays in Moral Philosophy*. Oxford: Oxford University Press, pp. 132-147.
- Fowler, Timothy and Zofia Stemplowska, 2015. The Asymmetry Objection Rides Again: On the Nature and Significance of Justificatory Disagreement. *Journal of Applied Philosophy*, 32(2), pp. 133-146.
- Frankfurt, Harry, 1988. Identification and Wholeheartedness. In: *The Importance of What We Care About*. New York: Cambridge University Press, pp. 159-176.
- Frazer, Elizabeth and Nicola Lacey, 1995. Politics and the Public in Rawls' Political Liberalism. *Political Studies*, 43(2), pp. 233-247.

- Freeman, Samuel, 2003. Congruence and the Good of Justice. In: *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press, pp. 277-315.
- Galeotti, Anna Elisabetta, 2002. *Toleration and Recognition*. Cambridge: Cambridge University Press.
- Gallie, W.B., 1956. Essentially Contested Concepts. *Proceedings of the Aristotelian Society*, Volume 56, pp. 167-198.
- Galston, William, 1995. Two Concepts of Liberalism. *Ethics*, 105(3), pp. 516-534.
- _____, 2010. Realism in Political Theory. *European Journal of Political Theory*, 9(4), pp. 385-411.
- Gaus, Gerald F., 1996. *Justificatory Liberalism. An Essay in Epistemology and Political Theory*. New York: Oxford University Press.
- _____, 1999. Reasonable Pluralism and the Domain of the Political: How the Weaknesses of John Rawls's Political Liberalism Can be Overcome by a Justificatory Liberalism. *Inquiry: An Interdisciplinary Journal of Philosophy*, 42(2), pp. 259-284.
- _____, 2011 a. *The Order of Public Reason. A Theory of Freedom and Morality in a Diverse and Bounded World*. New York: Oxford University Press.
- _____, 2011 b. A Tale of Two Sets: Public Reason in Equilibrium. *Public Affairs Quarterly*, 25(4), pp. 305-325.
- _____, 2015. On Dissing Public Reason: A Reply to Enoch. *Ethics*, 125(4), pp. 1078-1095.
- Geuss, Raymond, 2001. *History and Illusion in Politics*. Cambridge: Cambridge University Press.
- _____, 2008. *Philosophy and Real Politics*. Princeton (NJ): Princeton University Press.
- Gewirth, Alan, 1992. Human Dignity and the Basis of Rights. In: M. J. Meyer & W. Parent, eds. *The Constitution of Rights: Human Dignity and American Values*. Ithaca (NY): Cornell University Press, pp. 10-28.
- Gintis, Herbert, 2010. Social Norms as Choreography. *Politics, Philosophy & Economics*, 9(3), pp. 251-264.
- Goldman, Alvin I., 2010. Epistemic Relativism and Reasonable Disagreement. In: R. Feldman and T. A. Warfield, eds. *Disagreement*. Oxford: Oxford University Press, pp. 187-215.

- Goodin, Robert E. and Christian List, 2006. Conditional Defense of Plurality Rule : Generalizing May's Theorem in a Restricted Informational Environment. *American Journal of Political Science*, 50(4), pp. 940-949.
- Greene, Amanda, 2016. Consent and Political Legitimacy. In: D. Sobel, P. Vallentyne & S. Wall, eds. *Oxford Studies in Political Philosophy. Volume 2*. Oxford: Oxford University Press, pp. 71-96.
- _____, 2017. Legitimacy without Liberalism: A Defense of Max Weber's Standard of Political Legitimacy. *Analyse & Kritik*, 39(2), pp. 295-323.
- Habermas, Jürgen, 1996. *Between Facts and Norms. Contributions to a Discourse Theory of Law and Democracy*. Cambridge (Mass.): Polity Press.
- Hadfield, Gillian K. and Stephen Macedo, 2012. Rational Reasonableness: Toward a Positive Theory of Public Reason. *Law & Ethics of Human Rights*, 6(1), pp. 7-46.
- Hamlin, Alan and Zofia Stemplowska, 2012. Theory, Ideal Theory and the Theory of Ideals. *Political Studies Review*, 10(1), pp. 46-82.
- Hampton, Jean, 1989. Should Political Philosophy Be Done without Metaphysics?. *Ethics*, 99(4), pp. 791-814.
- Hart, H.L.A., 1955. Are There Any Natural Rights?. *The Philosophical Review*, 64(2), pp. 175-191.
- _____, 1958. Positivism and the Separation of Law and Morals. *Harvard Law Review*, 71(4), pp. 593-629.
- _____, 1961. *The Concept of Law*. Oxford: Clarendon Press.
- Hayek, Friedrich August, 1945. The Use of Knowledge in Society. *The American Economic Review*, 35(4), pp. 519-530.
- _____, 1989. The Pretence of Knowledge. *The American Economic Review*, 79(6), pp. 3-7.
- Himma, Kenneth Eimar, 2004. Inclusive Legal Positivism. In: J. L. Coleman, K. E. Himma and S. J. Shapiro, eds. *The Oxford Handbook of Jurisprudence and Philosophy of Law*. Oxford: Oxford University Press, pp. 125-165.
- Hinton, Timothy, 2001. Must Egalitarians Choose between Fairness and Respect?. *Philosophy & Public Affairs*, 30(1), pp. 72-87.
- Hodgson, Louis-Philippe, 2010. Kant on the Right to Freedom: A Defense. *Ethics*, 120(4), pp. 791-819.

- Honig, Bonnie, 1993. *Political Theory and the Displacement of Politics*. Ithaca (NY): Cornell University Press.
- Honneth, Axel, 2001. Invisibility: On the Epistemology of 'Recognition'. *Aristotelian Society Supplementary Volume*, 75(1), pp. 111-126.
- Hooker, Brad, 2012. Theory versus Anti-Theory in Ethics. In: U. Heuer and G. Lang, eds. *Luck, Value, and Commitment: Themes from the Ethics of Bernard Williams*. Oxford: Oxford University Press, pp. 19-40.
- Horton, John, 1998. Toleration As a Virtue. In: D. Heyd, ed. *Toleration: An Elusive Virtue*. Princeton (NJ): Princeton University Press, pp. 28-43.
- _____, 2012. Political Legitimacy, Justice, and Consent. *Critical Review of International Social & Political Philosophy*, 15(2), pp. 129-148.
- Hurka, Thomas, 1993. *Perfectionism*. Oxford : Oxford University Press.
- Hurley, Susan, 2003. *Justice, Luck, and Knowledge*. Cambridge (Mass.): Harvard University Press.
- Jones, Peter, 1994. Bearing the Consequences of Belief. *The Journal of Political Philosophy*, 2(1), pp. 24-43.
- _____, 2006. Toleration, Recognition and Identity. *The Journal of Political Philosophy*, 14(2), pp. 123-143.
- _____, 2015. Toleration, Religion, and Accommodation. *European Journal of Philosophy*, 23(3), pp. 542-563.
- Jubb, Robert, 2015. Playing Kant at the Court of King Arthur. *Political Studies*, 63(4), pp. 919-934.
- _____, 2017. Realism. In: A. Blau, ed. *Methods in Analytical Political Theory*. Cambridge: Cambridge University Press, pp. 112-130.
- Kant, Immanuel, [1785] 1996. Groundwork of the Metaphysics of Morals. In: M. J. Gregor, ed. *Practical Philosophy*. Cambridge: Cambridge University Press, pp. 43-108.
- _____, [1788] 1999. Critique of Practical Reason. In: M. J. Gregor, ed. *Practical Philosophy*, pp. 137-271.
- _____, [1797] 1996. The Metaphysics of Morals. In: M. J. Gregor, ed. *Practical philosophy*. New York: Cambridge University Press, pp. 353-603.

Kelly, Thomas, 2010. Peer-Disagreement and Higher-Order Evidence. In: R. Feldman & T. A. Warfield, eds. *Disagreement*, pp. 111-174.

_____, 2013. Disagreement and the Burdens of Judgement. In: D. Christensen & J. Lackey, eds. *The Epistemology of Disagreement: New Essays*. Oxford: Oxford University Press, pp. 31-52.

Kogelmann, Brian and Stephen J. Stich, 2016. When Public Reason Fails Us: Convergence Discourse as Blood Oath. *The American Political Science Review*, 110(4), pp. 717-730.

Kolodny, Niko, 2003. Love as Valuing a Relationship. *The Philosophical Review*, 112(2), pp. 135-189.

_____, 2014. Rule Over None I: What Justifies Democracy ?. *Philosophy & Public Affairs*, 42(3), pp. 195-229.

Kornhauser, Lewis A. and Lawrence G. Sager, 2004. The Many as One: Integrity and Group Choice in Paradoxical Cases. *Philosophy & Public Affairs*, 32(3), pp. 249-276.

Korsgaard, Christine, 1983. Two Distinctions in Goodness. *The Philosophical Review*, 92(2), pp. 169-195.

_____, 1986. Kant's Formula of Humanity. *Kant-Studien*, 77(2), pp. 183-202.

_____, 2013. The Relational Nature of the Good. In: R. Shafer-Landau, ed. *Oxford Studies in Metaethics. Volume 8*. Oxford: Oxford University Press, pp. 1-25.

Kukathas, Chandran, 1992. Are There Any Cultural Rights ?. *Political Theory*, 20(1), pp. 105-139.

_____, 2003. *The Liberal Archipelago. A Theory of Diversity and Freedom*. Oxford: Oxford University Press.

Kymlicka, Will, 1995. *Multicultural Citizenship: A Liberal Theory of Minority Rights*. Oxford: Oxford University Press.

_____, 2001. *Politics in the Vernacular: Nationalism, Multiculturalisms, and Citizenship*. Oxford: Oxford University Press.

_____, 2007. *Multicultural Odysseys. Navigating the New International Politics of Diversity*. New York: Oxford University Press.

Laborde, Cécile, 2015. Religion in the Law: The Disaggregation Approach. *Law & Philosophy*, 34(6), pp. 581-600.

_____, 2017. *Liberalism's Religion*. Cambridge (Mass.): Harvard University Press.

- Larmore, Charles, 1996. *The Morals of Modernity*. Cambridge: Cambridge University Press.
- _____, 1999. The Moral Basis of Political Liberalism. *The Journal of Philosophy*, 96(12), pp. 599-625.
- Lefkowitz, David, 2007. On a Moral Right to Civil Disobedience. *Ethics*, 117(2), pp. 202-233.
- Leiter, Brian, 2013. *Why Tolerate Religion?*. Princeton (NJ): Princeton University Press.
- Leland, R.J. and Han Van Wietmarschen, 2012. Reasonableness, Intellectual Modesty, and Reciprocity in Political Justification. *Ethics*, 122(4), pp. 721-747.
- Lenta, Patrick, 2016. Freedom of Conscience and the Value of Personal Integrity. *Ratio Juris*, 29(2), pp. 246-263.
- Levy, Jacob T., 2017. There is No Such Thing as an Ideal Theory. *Social Philosophy & Policy*, 33(1), pp. 312-333.
- List, Christian and Robert Goodin, 2001. Epistemic Democracy: Generalizing the Condorcet Jury Theorem. *The Journal of Political Philosophy*, 9(3), pp. 277-306.
- Lister, Andrew, 2010. Public Justification and the Limits of State Action. *Politics, Philosophy & Economics*, 9(2), pp. 151-175.
- _____. Public Reason and Reciprocity. *The Journal of Political Philosophy*, 25(2), pp. 155-172.
- Lyons, David, 1998. Moral Judgment, Historical Reality, and Civil Disobedience. *Philosophy & Public Affairs*, 27(1), pp. 31-49.
- Macedo, Stephen, 1997. In Defense of Liberal Public Reason: Are Slavery and Abortion Hard Cases?. *American Journal of Jurisprudence*, 42(1), pp. 1-29.
- Maclure, Jocelyn and Charles Taylor, 2011. *Secularism and Freedom of Conscience*. Cambridge (Mass.): Harvard University Press.
- Margalit, Avishai and Joseph Raz, 1990. National Self-Determination. *The Journal of Philosophy*, 87(9), pp. 439-461.
- Marmor, Andrei, 2015. What is the Right to Privacy?. *Philosophy & Public Affairs*, 43(1), pp. 3-26.
- May, Kenneth O., 1952. A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision. *Econometrica*, 20(4), pp. 680-684.
- May, Simon Căbulea, 2017. Exemptions for Conscience. In: C. Laborde & A. Bardon, eds. *Religion in Liberal Political Philosophy*, pp. 191-203.

- McMahon, Christopher, 2014. Rawls, Reciprocity and the Barely Reasonable. *Utilitas*, 26(1), pp. 1-22.
- Meckled-Garcia, Saladin, 2016. Two Concepts of Justice - and of Its Scope. *Critical Review of International Social & Political Philosophy*, 19(5), pp. 534-554.
- Mill, John Stuart, [1859] 2015. On Liberty. In: M. Philp & F. Rosen, eds. *On Liberty, Utilitarianism, and Other Essays*. New York: Oxford University Press, pp. 1-112.
- Miller, David, 2012. Territorial Rights: Concept and Justification. *Political Studies*, 60(2), pp. 252-268.
- _____, 2013. Political Philosophy for Earthlings. In: *Justice for Earthlings. Essays in Political Philosophy*. Cambridge: Cambridge University Press, pp. 16-39.
- Mills, Charles W., 2017. 'Ideal Theory' as Ideology. In: *Black Rights/White Wrongs: The Critique of Racial Liberalism*. Oxford: Oxford University Press, pp. 72-90.
- Moore, George Edward, 1922. The Conception of Intrinsic Value. In: *Philosophical Studies*. London: Routledge and Kegan Paul.
- Nagel, Thomas, 1987. Moral Conflict and Political Legitimacy. *Philosophy & Public Affairs*, 16(3), pp. 215-240.
- _____, 1998. Concealment and Exposure. *Philosophy & Public Affairs*, 27(1), pp. 3-30.
- Nine, Cara, 2008. A Lockean Theory of Territory. *Political Studies*, 56(1), pp. 148-165.
- Nozick, Robert, [1974] 2013. *Anarchy, State and Utopia*. New York: Basic Books.
- Nussbaum, Martha, 2008. *Liberty of Conscience*. New York: Basic Books.
- _____, 2011. Perfectionist Liberalism and Political Liberalism. *Philosophy & Public Affairs*, 39(1), pp. 3-45.
- _____, 2015. Political Liberalism and Global Justice. *Journal of Global Ethics*, 11(1), pp. 68-79.
- O'Neill, Onora, 1987. Abstraction, Idealization and Ideology in Ethics. *Royal Institute of Philosophy Supplement*, Volume 22, pp. 55-69.
- Ostrom, Elinor, 1990. *Governing the Commons. The Evolution of Institutions of Collective Action*. Cambridge: Cambridge University Press.
- Ottonelli, Valeria, 2012. Equal Respect, Equal Competence and Democratic Legitimacy. *Critical Review of International Social & Political Philosophy*, 15(2), pp. 201-218.

Pallikkathayil, Japa, 2010. Deriving Morality from Politics: Rethinking the Formula of Humanity. *Ethics*, 121(1), pp. 116-147.

_____, 2016. Neither Perfectionism nor Political Liberalism. *Philosophy & Public Affairs*, 44(3), pp. 171-196.

Patten, Alan, 2014. *Equal Recognition. The Moral Foundations of Minority Rights*. Princeton (NJ): Princeton University Press.

Peter, Fabienne, 2013. Epistemic Foundations of Political Liberalism. *Journal of Moral Philosophy*, 10(5), pp. 598-620.

Pettit, Philip, 2001. Deliberative Democracy and the Discursive Dilemma. *Philosophical Issues*, 11(1), pp. 268-299.

Philpott, Daniel, 2016. *Sovereignty*, *The Stanford Encyclopaedia of Philosophy*. [Online] Available at: <https://plato.stanford.edu/archives/sum2016/entries/sovereignty/> [Accessed 20 08 2018].

Quong, Jonathan, 2006. Cultural Exemptions, Expensive Tastes, and Equal Opportunities. *Journal of Applied Philosophy*, 23(1), pp. 53-71.

_____, 2011. *Liberalism without Perfection*. Oxford: Oxford University Press.

Rawls, John, 1999 [1951]. Outline of a Decision Procedure for Ethics. In: S. Freeman, ed. *Collected Papers*. Cambridge (Mass.): Harvard University Press, pp. 1-19.

Rawls, J., 1999 [1971]. *A Theory of Justice. Revised Edition*. Cambridge (Mass.): Belknap Press of Harvard University Press.

_____, 1999 [1985]. Justice as Fairness: Political not Metaphysical. In: S. Freeman, ed. *Collected Papers*. Cambridge (Mass.): Harvard University Press, pp. 388-414.

_____, 1999 [1987]. The Idea of an Overlapping Consensus. In: S. Freeman, ed. *Collected Papers*. Cambridge (Mass.): Harvard University Press, pp. 421-448.

_____, 1999. *The Law of Peoples*. Cambridge (Massachusetts): Harvard University Press.

_____, 2001. *Justice as Fairness. A Restatement*. Cambridge, Mass.: Belknap Press of Harvard University Press.

_____, 2005 [1993]. *Political Liberalism. Expanded Edition*. New York : Columbia University Press.

_____, 2005 [1997]. The Idea of Public Reason Revisited. In: *Political Liberalism*. New York: Columbia University Press, pp. 440-490.

- Raz, Joseph, 1986. *The Morality of Freedom*. Oxford : Clarendon Press.
- _____, 1995 a. Multiculturalism: A Liberal Perspective. In: *Ethics in the Public Domain: Essays in the Morality of Law and Politics*. Oxford: Oxford University Press, pp. 170-191.
- _____, 1995 b. Free Expression and Personal Identification. In: *Ethics in the Public Domain: Essays in the Morality of Law and Politics*. Oxford: Oxford University Press, pp. 146-169.
- _____, 1998. Disagreement in Politics. *American Journal of Jurisprudence*, 43(1), pp. 25-52.
- _____, 2001. *Value, Respect, and Attachment*. New York: Cambridge University Press.
- _____, 2010. Human Rights without Foundations. In: S. Besson & J. Tasioulas, eds. *The Philosophy of International Law*. Oxford: Oxford University Press, pp. 378-399.
- Ripstein, Arthur, 2004. Authority and Coercion. *Philosophy & Public Affairs*, 32(1), pp. 2-35.
- Rorty, Richard, 1989. *Contingency, Irony, and Solidarity*. New York: Cambridge University Press.
- Ross, William David, 2002 [1930]. *The Right and the Good*. Oxford: Oxford University Press.
- Rossi, Enzo, 2012. Justice , Legitimacy and (Normative) Authority for Political Realists. *Critical Review of International Social & Political Philosophy*, 15(2), pp. 149-164.
- Rossi, Enzo and Matt Sleat, 2014. Realism in Normative Political Theory. *Philosophy Compass*, 9(10), pp. 689-701.
- Satz, Debra, 2007. Equality, Adequacy, and Education for Citizenship. *Ethics*, 117(4), pp. 623-648.
- Scanlon, Thomas, 1998. *What We Owe to Each Other*. Cambridge (Mass.): Belknap Press of Harvard University Press.
- _____, 2003 a. Contractualism and Utilitarianism. In: *The Difficulty of Tolerance: Essays in Political Philosophy*. New York: Cambridge University Press, pp. 124-150.
- _____, 2003 b. A Theory of Freedom of Expression. In: *The Difficulty of Tolerance: Essays in Political Philosophy*. Cambridge: Cambridge University Press, pp. 6-25.
- _____, 2003 c. The Difficulty of Tolerance. In: *The Difficulty of Tolerance*. Cambridge: Cambridge University Press, pp. 187-201.
- _____, 2003 d. The Diversity of Objections to Inequality. In: *The Difficulty of Tolerance: Essays in Political Philosophy*. Cambridge (Mass.): Harvard University Press, pp. 202-217.

- ____, 2006. Reasons: A Puzzling Duality?. In: R. J. Wallace, P. Pettit, S. Scheffler & M. Smith, eds. *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. Oxford: Clarendon Press, pp. 231-246.
- Scheffler, Samuel, 1994. The Appeal of Political Liberalism. *Ethics*, 105(1), pp. 4-22.
- ____, 2003. What is Egalitarianism?. *Philosophy & Public Affairs*, 31(1), pp. 5-39.
- ____, 2007. Immigration and the Significance of Culture. *Philosophy & Public Affairs*, 35(2), pp. 93-125.
- Schmidtz, David, 2011. Nonideal Theory: What It Is and What It Needs to Be. *Ethics*, 121(4), pp. 772-796.
- Schoenfield, Miriam, 2014. Permission to Believe: Why Permissivism is True and What It Tells Us About Irrelevant Influences on Belief. *Noûs*, 48(2), pp. 193-218.
- Schwartzman, Micah, 2011. The Sincerity of Public Reason. *The Journal of Political Philosophy*, 19(4), pp. 375-398.
- ____, 2012. What If Religion Is Not Special?. *The University of Chicago Law Review*, 79(4), pp. 1351-1472.
- Sen, Amartya, 2006. What Do We Want from a Theory of Justice?. *The Journal of Philosophy*, 103(5), pp. 215-238.
- ____, 2009. *The Idea of Justice*. Cambridge (Mass.): Harvard University Press.
- Shelby, Tommie, 2007. Justice, Deviance, and the Dark Ghetto. *Philosophy & Public Affairs*, 35(2), pp. 126-160.
- Sher, George, 1997. *Beyond Neutrality. Perfectionism and Politics*. New York: Cambridge University Press.
- Sidgwick, Henry, 1884. *The Methods of Ethics*. Third Edition ed. London: Macmillan and Co.
- Simmons, A. John, 1979. *Moral Principles and Political Obligations*. Princeton (NJ): Princeton University Press.
- ____, 1987. The Anarchist Position: A Reply to Klosko and Senor. *Philosophy & Public Affairs*, 16(3), pp. 269-279.
- ____, 1999. Justification and Legitimacy. *Ethics*, 109(4), pp. 739-771.
- ____, 2010. Ideal and Nonideal Theory. *Philosophy & Public Affairs*, 38(1), pp. 5-36.

- _____, 2015. Territorial Rights. In: D. Sobel, P. Vallentyne & S. Wall, eds. *Oxford Studies in Political Philosophy, Volume 1*, pp. 145-171.
- Simpson, Robert Mark, 2013. Epistemic Peerhood and the Epistemology of Disagreement. *Philosophical Studies*, 164(2), pp. 561-577.
- Sinclair, Thomas Porter, 2018. The Power of Public Positions. In: D. Sobel, P. Vallentyne & S. Wall, eds. *Oxford Studies in Political Philosophy. Volume 4*. Oxford: Oxford University Press, pp. 28-52.
- Sleat, Matt, 2014. Legitimacy in Realist Thought: Between Moralism and Realpolitik. *Political Theory*, 42(3), pp. 314-337.
- Southwood, Nicholas, 2011. The Moral/Conventional Distinction. *Mind*, 120(479), pp. 761-802.
- Stilz, Anna, 2011. Nations, States, and Territories. *Ethics*, 121(3), pp. 572-602.
- Strawson, Peter Frederick, 1974. Freedom and Resentment. In: *Freedom and Resentment, and Other Essays*. London: Meuthen, pp. 1-25.
- Sunstein, Cass R. and Richard Thaler, 2003. Libertarian Paternalism Is Not an Oxymoron. *The University of Chicago Law Review*, 70(4), pp. 1159-1202.
- Taylor, Charles, 1994. The Politics of Recognition. In: *Multiculturalism*. Princeton (NJ): Princeton University Press, pp. 25-73.
- Thrasher, John and Kevin Vallier, 2015. The Fragility of Consensus: Public Reason, Diversity and Stability. *European Journal of Philosophy*, 23(4), pp. 933-954.
- Valentini, Laura, 2012. Justice, Disagreement and Democracy. *British Journal of Political Science*, 43(1), pp. 177-199.
- _____, 2017. Dignity and Human Rights: A Reconceptualisation. *Oxford Journal of Human Rights*, 37(4), pp. 862-885.
- Vallier, Kevin, 2011. Against Public Reason Liberalism's Accessibility Requirement. *Journal of Moral Philosophy*, 8(3), pp. 366-389.
- _____, 2012. Liberalism, Religion and Integrity. *Australasian Journal of Philosophy*, 90(1), pp. 149-165.
- _____, 2016 a. In Defence of Intelligible Reasons in Public Justification. *The Philosophical Quarterly*, 66(264), pp. 596-616.
- _____, 2016 b. The Moral Basis of Religious Exemptions. *Law & Philosophy*, 35(1), pp. 1-28.

- _____, 2018. "Public Justification", *The Stanford Encyclopaedia of Philosophy*. [Online] Available at: <<https://plato.stanford.edu/archives/spr2018/entries/justification-public/>> [Accessed 11 09 2018].
- Van Schoelandt, Chad, 2015. Justification, Coercion, and the Place of Public Reason. *Philosophical Studies*, 172(4), pp. 1031-1050.
- Van Wietmarschen, Han, 2013. Peer-Disagreement, Evidence, and Well-Foundedness. *The Philosophical Review*, 122(3), pp. 395-425.
- _____, forthcoming. Reasonable Citizens and Epistemic Peers: A Skeptical Problem for Political Liberalism. *The Journal of Political Philosophy*.
- Waldron, Jeremy, 1987. Theoretical Foundations of Liberalism. *The Philosophical Quarterly*, 37(147), pp. 127-150.
- _____, 1992. Minority Cultures and the Cosmopolitan Alternative. *University of Michigan Journal of Law Reform*, Volume 25, pp. 751-794.
- _____, 1996. Kant's Legal Positivism. *Harvard Law Review*, 109(7), pp. 1535-1566.
- _____, 1999. *Law and Disagreement*. Oxford: Oxford University Press.
- _____, 2006. The Core of the Case Against Judicial Review. *The Yale Law Journal*, 115(6), pp. 1346-1406.
- Wall, Stephen, 1998. *Liberalism, Perfectionism and Restraint*. Cambridge: Cambridge University Press.
- Weithman, Paul, 2011. *Why Political Liberalism? On Rawls's Political Turn*. New York: Oxford University Press.
- _____, 2016. Inclusivism, Stability and Assurance. In: *Rawls, Political Liberalism and Reasonable Faith*. New York: Cambridge University Press, pp. 150-167.
- White, Roger, 2005. Epistemic Permissiveness. *Philosophical Perspectives*, 19(1), pp. 445-459.
- Williams, Bernard, 1980. A Critique of Utilitarianism. In: *Utilitarianism: For and Against*. Cambridge: Cambridge University Press, pp. 75-150.
- _____, 1981 a. Internal and External Reasons. In: *Moral Luck*. Cambridge: Cambridge University Press.
- _____, 1981 b. Persons, Character and Morality. In: *Moral Luck*. Cambridge: Cambridge University Press, pp. 1-19.

___, 1998. Toleration: An Impossible Virtue?. In: D. Heyd, ed. *Toleration: An Elusive Virtue*, pp. 18-27.

___, 2005. Realism and Moralism in Political Theory. In: *In the Beginning was the Deed*. Princeton (NJ): Princeton University Press, pp. 1-17.

Wolff, Jonathan, 1998. Fairness, Respect, and the Egalitarian Ethos. *Philosophy & Public Affairs*, 27(2), pp. 97-122.

___, 2017. Forms of Differential Social Inclusion. *Social Philosophy & Policy*, 34(1), pp. 164-185.

Wolff, Robert Paul, 1998 [1970]. *In Defense of Anarchism*. Los Angeles: University of California Press.

Young, Iris Marion, 1990. *Justice and the Politics of Difference*. Princeton (NJ): Princeton University Press.

Ypi, Lea, 2014. A Permissive Theory of Territorial Rights. *European Journal of Philosophy*, 22(2), pp. 288-312.