

Against Justificatory Liberalism

B.Phil. Thesis

13 June 2012

Matthias Brinkmann

matthias.brinkmann@philosophy.ox.ac.uk

(Original version with reformat for readability.)

Abstract

Against Justificatory Liberalism

In this thesis I argue that we should give up *justificatory liberalism*: the claim that it is a necessary condition for the legitimacy of a law that it is publicly justifiable.

My argument proceeds in three stages. First, I give an intuitive motivation for justificatory liberalism in the introduction (ch. 1). Then, I sketch the most plausible version of justificatory liberalism, and distinguish different forms of it (ch. 2). I analyse the central notion of “public justification”, which I interpret through the idea that all citizens *could reasonably accept* a law. What citizens could reasonably accept is a mix of moral, epistemic and counterfactual considerations, and is constrained by internalism. I also argue that justificatory liberalism is analytically independent from claims about justificatory neutrality and duties of civility.

In the second stage (chs. 3 & 4), I argue that justificatory liberalism falls prey to the *Basic Problem*. There is reasonable disagreement about virtually all laws, policies, and institutional arrangements. But if we accept justificatory liberalism, then few if any non-trivial laws will be legitimate. This contradicts a liberal commitment to anti-anarchism, according to which many such laws are legitimate. I then discuss a number of solutions to the Basic Problem that justificatory liberals have proposed (ch. 4): actual agreements, procedural meta-agreement, the special nature of the political, moralised agreement, and comparative agreement. While these solution attempts are sophisticated, they all fail.

I discuss the implications of this result in the third stage (ch. 5). In reply to the Basic Problem trilemma, we must ask which liberal commitment we ought to give up – justificatory liberalism, reasonable disagreement, or anti-anarchism. I argue on principled grounds that the latter two are central (moral) principles of liberal thought, while the former is not. In particular, equal respect – the core notion in all liberal views – does not require justificatory liberalism. Thus, we should give up justificatory liberalism.

This argument leaves open many interesting questions that surround justificatory liberalism. I outline some questions for further research in my concluding chapter (ch. 6).

Contents

1	Introduction	4
2	Justificatory Liberalism	6
2.1	Coercion	6
2.2	Legitimacy	7
2.3	Public Justification	9
2.4	“Could reasonably accept”	11
2.5	Acceptance and Rejection	16
2.6	Refinements	19
2.7	Connections to Other Debates	21
3	The Basic Problem	25
3.1	Reasonable Disagreement	25
3.2	Anti-Anarchism	26
3.3	The Debate So Far	27
3.4	Reflexivity	28
3.5	Baselines	33
4	Failed Solutions of the Basic Problem	35
4.1	Actual Agreements	35
4.2	Procedural Meta-Agreement	37
4.3	Asymmetries	39
4.4	Moral Content	43
4.5	Comparisons	49
5	The Weakest Link	56
5.1	Liberalism and Respect	56
5.2	Respect and Justification	60
5.3	Other Arguments	68
5.4	Liberal Commitments: Reasonable Disagreement	71
5.5	Liberal Commitments: Anti-Anarchism	72
6	Concluding Remarks	74
	Works Cited	76

1 Introduction

One of the most basic problems in political philosophy is that of coercion. Individuals live together in states, which exercise power over them: they thwart individuals' aims, submit them to an authority they often oppose, and interfere in many other ways with people's lives. For these reasons, coercion is generally held to be *prima facie* wrong. Strong moral reasons are needed to overcome the liberal presumption against coercion.

Social contract theorists have been very impressed by this. They try to overcome the *prima facie* wrongness of coercion by showing that individuals *do or can consent* to the state's exercise of coercive power. Consent is normatively significant: if I consent to the actions of a political authority, they (partially) lose their problematic character. We might even argue that consent is strong enough to make otherwise coercive acts *non-coercive*.

Alternatively, the problem and the solution can be explained in Rousseauian terms: alienation between individual and general will is the problem that unsettles liberals. The liberal aim therefore is to find a way how the individual will can be reconciled with the general will. The most obvious reconciliation is through the individual's actual or hypothetical consent.¹

Most modern liberals have moved away from the narrow notion of consent. The wider idea they pursue is that a law has to be *justified to every citizen*, or *publicly justified*. Public justification is more than rational or moral justification *simpliciter*; it is justification *to people*. Successful public justification shows that all citizens have a reason to accept certain state actions, and that they could reasonably accept such actions.

Despite these changes, the basic project remains the same. What is publicly justified primarily depends on the *actual* interests and beliefs of citizens, even if actual consent is no longer held to be solely relevant. Thus, modern liberals insist, the public justification of a law is normatively significant, and if public justification of a law is achieved, this lessens or entirely eliminates its coercive and alienating character.

This suggests a first claim,

If a coercive law is justified to all citizens, then it is legitimate.

This simple claim needs to be amended. For example, laws given by usurpers or through unjust procedures might fail to be legitimate, even if they are publicly justified as far as their content is concerned.

¹ See Manin 1987; Dryzek and Niemeyer 2006.

However, in this thesis I want to consider the more controversial claim that public justification is a *necessary* condition on legitimacy. This position I call

Justificatory Liberalism (JL):² If a coercive law is legitimate, then it is publicly justifiable.

There are likely to be further necessary conditions on the exercise of legitimate power, but I will not be concerned with such conditions here.

(JL) has first been made prominent by John Rawls through his “liberal principle of legitimacy”:

“our exercise of political power is fully proper³ 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.” (2005, 137)

Since Rawls’ original formulation, many political philosophers have accepted a version of (JL) – to the degree that one might call it the majority position.⁴ However, I will argue that we should reject it. My critique will be an internal critique: I will argue that, *as liberals*, we should reject (JL).

My argument proceeds in two stages:

Basic Problem Argument. (JL) is incompatible with two other central liberal commitments, reasonable disagreement and anti-anarchism.

Weakest Link Argument. If liberals have to give up one of these three commitments, it should be (JL).

Thus, liberals should give up (JL).

I describe justificatory liberalism in greater detail in section 2, before I outline the “Basic Problem” in section 3. Section 4 deals with various attempts to avoid the Basic Problem that have been offered by justificatory liberals. I argue that they all fail. Given that result, I proceed to claim in section 5 that (JL) is the “weakest link” among the three premises that give rise to the Basic Problem, and should thus be given up.

² I take this label from Gaus 1996; Eberle 2002.

³ “Fully proper” is different from “legitimate”. Possibly, Rawls shied away from stating (JL) here (cf. Copp 2011, 258–9). In general, I will be unconcerned with Rawls exegesis in this thesis, and will use themes from Rawls liberally (cf. sec. 3.3).

⁴ For some better-known advocates, see Nagel 1987, Larmore 1990, Gaus 1996, Klosko 2003, Estlund 2008, Quong 2011, and many others.

2 Justificatory Liberalism

To gain a good understanding of (JL), it is best to analyse the terms used in it:

If a coercive law is legitimate, then it is publicly justifiable.

I will discuss the three underlined terms now in turn.

2.1 Coercion

First, (JL) has to specify *which laws*⁵ are subject to a requirement of public justification. Call this the *scope* of (JL).⁶ The formulation I used in the introduction suggests that (JL)'s scope are *coercive* laws. The justification of coercion has been a classical concern of liberals such as Rawls, Nagel and Dworkin.

Paradigmatic cases of coercion on a common-language understanding of that word are threats and using physical force. However, not all government policies are coercive in this sense (Morris 2012). Governments issue health recommendations, regulate the (voluntary) institution of marriage, and adjust inflation levels. But this gives some wrong results (Bird 2011). For example, if a government decides to display religious symbols in public, this is hardly coercive. But justificatory liberals commonly regard such displays as a paradigmatic policy which requires public justification.

An alternative approach to (JL)'s scope would be to focus on law as a set of *rules* that claim *authority*. Here, coercion is seen as secondary to the nature of law.⁷ As virtually all government policies claim to be authoritative, this criterion would result in a very wide scope. On the other hand, this criterion fails to highlight the specifically problematic nature of coercion; there is a difference in normative significance between government health recommendations and the rules of criminal law – the latter, we think, stand under a more stringent requirement of public justification than the former.

The difference between these two scope criteria might look negligible, but they correspond to two different ways of stating the problem of legitimacy, which already surfaced in the introduction. We can see (JL) as a normative response to the problem of coercion or to the problem of reconciling individual will with general will (or authority). These two ways of setting out the problem incline us to the first and second scope criteria, respectively.

⁵ I will interchangeably speak of "policies" throughout this thesis.

⁶ "Scope" (and "constituency" below) are labels adopted from Quong 2011, ch. 9.

⁷ See Ripstein 2004 for this distinction.

The first way of motivating (JL) is by far more common,⁸ and so I will focus on coercion-based justificatory liberalism in this thesis. Still, it is methodologically helpful to consider forms of (JL) with a wide, if not universal, scope. This can be achieved by moving to a Kantian meaning of “coercion”. For Kantians, coercion is widened beyond its ordinary sense, and understood to mean any “hindrance” or restriction of an individual’s freedom (Ripstein 2004). Such an expansive meaning is how most liberals probably mean to understand the term anyway. Throughout this thesis, then, coercion does not connote the narrow sense of coercion as “bolts and shackles”, but any interference of the government with an individual’s life. A further question about scope is raised in the next section.

Independent from the scope of (JL), we need to specify *towards whom* justification is due. Call this the *constituency* of (JL). A natural reading would be that justification is due to all adult citizens who have minimal reasoning capacities; but other criteria are imaginable – maybe justification is due to all citizens whose lives are influenced by the policy in question. These are difficult and well-known questions,⁹ on which I will remain silent here (with one exception in sec. 2.4.4). I will use the placeholder “citizens” to denote the group of persons to whom justification is due. The reader should be aware that I use “citizens” in this technical (and artificial) sense throughout this thesis.

2.2 Legitimacy

We can now turn to legitimacy. Like few other concepts, legitimacy is an “essentially contested concept”. I can not defend my usage here, nor do I intend to decide what the best analysis of that concept is. Nevertheless, by *legitimacy* I understand the moral *permissibility* of the state to issue certain laws, and to coerce citizens on the basis of those laws. Importantly, I will assume that this does *not* imply the existence of an individual duty to obey, or a right to rule on part of the legitimate authority.

Legitimacy, as relating to the deontic category of the permissible, describes a *range* of options, and is a *binary property*. As such it contrasts with justice, which describes a *ranking* of options, and is a *gradual property*. We evaluate the justice of laws and institutions on the basis of certain positive traits they possess. Insofar as some possess these traits to a higher degree, we speak of them as being “more just” than others. It makes little sense, however, to speak of “more” or “less” legitimate exercises

⁸ See sec. 5.4.

⁹ Compare the “boundary problem” in democratic theory – e.g., Goodin 2007.

of coercion, as it makes no sense to speak of “more” or “less” permissible actions.¹⁰

Questions about legitimacy can be asked on *different levels*. On the most general level, we can ask whether a certain government or form of government is legitimate. On a medium level, we focus on the legitimacy of constitutions and the “basic structure” of a society, and the forms of coercion entailed by it. On the lowest level, we look to particular coercive act-tokens of particular government agents – police officers, tax collectors, etc. – and judge the legitimacy of these acts.

It is tempting to reduce all facts about legitimacy to one level. On the micro-reductive approach, facts about the legitimacy of particular coercive act-tokens are the primary facts; on the macro-reductive approach, we take facts about the legitimacy of general laws, or the institutions of a state, as most basic.

But neither of these approaches is correct. First, against the micro-reductive approach, sometimes the procedural origin of certain coercive laws makes them legitimate, even if they would be illegitimate otherwise. Some taxation scheme might fail to be legitimate under a non-democratic regime; but a law with identical content might be legitimate if decided through a legitimate decision procedure.

Second, I reject the extreme proceduralism found in a macro-reductive approach. Legitimate decision procedures do *not* invariably bestow legitimacy on particular laws and coercive act-tokens. Similarly, legitimate *general* laws or constitutions do not guarantee the legitimacy of more *particular* laws which are connected to them.

Thus, there are irreducible but connected facts about legitimacy on different levels of generality. In this thesis I want to set aside questions about the legitimacy of the institutions and procedures through which law is made, and the legitimacy of a given state or government. I will focus on the normal case, in which laws and policies are decided through a legitimate decision procedure by a legitimate government in a legitimate state. This does not deny the relevance of legitimacy questions on these levels, but is intended to simplify the discussion to a case we easier tackle.

Laws and policies in turn come on different levels of generality. Some belong to (written or unwritten) constitutions, and determine the “basic structure” of a society; others regulate highly specific aspects of public life. Some justificatory liberals like Rawls have wanted to limit the scope of (JL) to laws of high generality, the “constitutional essentials”. Such a scope-restricted justificatory liberalism is less demanding than more expansive

¹⁰ We often speak of more and less legitimate *authorities*. We can interpret this to mean that the more legitimate authority is permitted a wider set of coercive actions than the less legitimate authority.

forms, and we might think that it avoids many of the problems that I will outline later.

Many commentators have argued that the distinction between “constitutional essentials” and other laws is too vague, or can not be upheld.¹¹ But even if we can draw this distinction clearly, such a scope restriction contradicts the motivation that underlies justificatory liberalism. On the coercion-focused version I outlined, we are committed to justify interferences with individuals’ lives. But many of the areas of law which are problematic in this sense are commonly regulated on a sub-constitutional level – e.g., levels regulating abortion, freedom of speech, punishment for crimes, or the social welfare net.

Unquestionably, the constitutional content is usually indirectly relevant to these more particular questions, insofar as a constitution permits or forbids these more particular laws. But by and large, the constitutional—non-constitutional distinction does not track the justification-demanding—non-justification-demanding distinction. Thus, I will assume that if a law can be shown to be coercive in the sense sketched in sec. 2.2, it will fall inside the scope of (JL), no matter what its level of generality.

2.3 Public Justification

I will now turn to the notion that a law is *publicly justifiable*. In a first sense of “public”, public justification is opposed to *secret* justification (Gosseries 2010). This would make (JL) a claim about how governments ought to communicate with citizens; and we would rule out certain forms of Sidgwickian “government house” utilitarianism.

A stronger sense of “public” is outlined by the early Rawls (Rawls 1999b [1971]; cf. Williams 1998). Social rules are counted as public by Rawls if they are known by everyone, and everyone knows that everyone knows these rules etc. Furthermore, public rules are sufficiently determinate – they allow each individual to infer what is permitted and forbidden – and individuals are assured of their mutual willingness to abide by these rules, given that others cooperate.

Liberals, whether justificatory or non-justificatory, will usually require that legitimate coercive laws are public in these two senses – that is, there should be some non-secret justification for them, and they should be widely known and followed (or, they should be widely *knowable* and *followable*).

Justificatory liberalism, however, proposes a more substantial reading of public justification. In a sense, we can read justificatory liberalism to start from the uncontroversial core idea of publicity and develop it into a stronger notion.¹² The justifica-

¹¹ Greenawalt 1993; Arneson 2003; Bird 2011. Cf. Wall 1996, 41–2.

¹² See Freeman 2007a for such an argument in the context of Rawls’s position.

tion of a coercive law is public in this stronger sense if that law is justifiable *to every* citizen.

There are different ways to cash out the idea of “justifiability to”. “Justifiability” suggests an essentially epistemic notion that focusses on standards of rationality, and people’s beliefs. The language of justification, however, easily gives the wrong impression that we tackle a primarily *epistemic* problem. But we should remember the origin of justificatory liberalism in consent theories: what bears normative significance is the individual *will*, the *voluntaristic attitude* that citizens have towards the laws they are subject to. The aim of justificatory liberals is not to bring citizens to *believe* the correctness of some moral theory; it is to show that some coercion is *acceptable*, and could be *endorsed*. (Believing coercion to be right is of course one important aspect of accepting it.)

Thus, we do better if we move to well-known language from existing contractualist theories about what people could *accept*, or could *reasonably accept*. From here on I will understand the phrase that some law *L* is justifiable to *A* to mean that *A* could reasonably accept *L*, a notion I will explain in the next subsection.

Before turning to that discussion, it is fruitful to ask how the *justifiability* of a law connects to its *justifiedness*. A law is publicly *justified* when we have actually *exercised* public justification, where public justification is understood to be a *process* of public argument, where we engage others, and present them with reasons for a proposed law. A law is publicly *justifiable*, on a first pass, if it is possible to successfully justify it under actual circumstances.

We should observe *first* that there are high epistemic barriers to determine what is publicly justifiable. It is hard to say what individuals could reasonably accept, as this requires us to apply a demanding counterfactual standard. Furthermore, we are often biased when we abstractly try to find out what others could accept: we tend to think that others could accept our preferred theory or solution. Actual public justification can play an important epistemic and evidential role: it helps us to determine what is publicly justifiable.

Second, it might be the case that individuals *could not* reasonably accept laws because they *do not actually* accept them – or phrased in the language of justification, there might be laws which are not publicly justifiable *because* they are not publicly justified. This sounds paradoxical, but is a natural claim. To take an analogous example from morality, we could never reasonably accept to be raped *against our actual consent*, no matter how unreasonable our actual consent (Parfit 2011, 191–200). The fact that the same sexual act would have been consensual had the participants been more reasonable bears no moral significance. It is unlikely, however, that there is a universal connection be-

tween actual acceptance and reasonable acceptability. Many policies are reasonably acceptable to idealised contractors even if their actual counterparts do not accept them and there is no successful public justification in favour of those policies.

Third, it will often be the case that actual public justification plays a transformative role on citizens: through *attempting* to publicly justify a law, it *becomes* publicly justifiable, as individuals' preferences and beliefs change and converge through discussion. Deliberative democrats often put huge emphasis on this transformative role.

Through these three connections, justificatory liberals will tend to assign actual public justification a central role in politics. However, the three connections are contingent and lack universality, and so we do not need to be concerned with actual public justification here.

2.4 “Could reasonably accept”

The language of what people can (or could) reasonably accept, or reasonably reject, has been made prominent through Thomas Scanlon's contractualism in moral philosophy (Scanlon 1998), and has been adopted by other moral philosophers (e.g., Parfit 2011).

The notion of what “people could reasonably accept” contains three different notions: (1) a *modal* element about what *would be the case* in certain counter-factual situations; (2) a *reason-based* claim on what it is *reasonable* for persons to accept; this splits into (2a) a moral sense, and (2b) an epistemic sense of “reasonable”; and (3) a specific *voluntaristic attitude* (“acceptance”) that is held to be relevant.

These elements could be worked out at great length, but a sketch must suffice here. I will start by describing the guiding idea behind the notion, before I turn to some details, and an advantage of this position over an alternative, Rawlsian approach.

2.4.1 The Internalism Constraint

Remember that we are interested in overcoming a presumption against coercion. Our recipe for overcoming such a presumption is to show that individuals have reason to accept that coercion – *actual* individuals, that is. We will impose idealisations on individuals to determine what they could reasonably accept. *But the stronger these idealisations are, the more we move away from what actual individuals have reason to accept.* This inverse relationship is hugely important. Put differently, the normative significance of what people could reasonably accept decreases with the thickness of modal abstraction that goes into the notions of “could” and “reasonable”.

Given this fact, all viable interpretations of “could reasonably accept” are constrained by *internalism*: to a large degree, what people could reasonably accept must be determined by connecting it to people’s actual beliefs, preferences, and desires. We must reject strong externalism, according to which what persons could reasonably accept are just the propositions which are true, or what we would believe if we were extremely reasonable. This is not due to an epistemic analysis of “could reasonably accept”,¹³ but due to an essentially *moral* concern (cf. sec. 5.4).

“Internalism” and “externalism” denote extremes on a scale. For example, what Gaus calls a “weak” form of externalism is compatible with the current position. According to this position, *A* could reasonably accept *p* if it is counterfactually true that in a process of “open” discussion, given new evidence, *A* could or would revise his beliefs in such a way that he believes that *p* (Gaus 1996, 137–43; 2011, ch. 13). This still takes the individual’s beliefs seriously to a sufficient degree.

One might be tempted to rephrase the current approach in terms of “deeply held” and “superficial” beliefs of individuals (cf. Raz 1998, 39–40). (Instead of “beliefs”, we could also write “identities”, “wants” or “preferences”.) The internalism constraint, under this interpretation, is tantamount to the claim that we always have to take into account the deeply held beliefs of citizens, but that we can disregard their more superficial beliefs. For example, claiming that a Catholic could reasonably accept atheism ignores some of her central commitments; on the other hand, we might be allowed to disregard some of her particular interpretations of Catholic doctrine.

However, some individuals, even if we idealise them only slightly, might have radically different core beliefs; and as I argue below, we can ascribe a basic sense of moral reasonableness to individuals, which will drastically change the beliefs of (e.g.) racists and anarchists. Still, the distinction between deeply held and superficial beliefs can often serve as a good but fallible heuristic.

2.4.2 “Could accept”

The best semantic analysis we should give of “could” in the current context is in terms of a might-counterfactual: *A* could reasonably accept *p* iff, were *A* in a counterfactual situation *S*, she might accept *p*.

This interpretation has two advantages: first, it makes it possible that *A* could reasonably accept *p*, while *A* could also reasonably accept not-*p*. Think back to our Catholic, and the various possible interpretations of Catholic doctrine. As an example, we want to be able to say that she could reasonably accept laws for abortion (under one interpretation of Catholic doctrine) and that

¹³ As Gaus seems to think; see Gaus 1996, 32–4, 140–1; 2003, 150–2.

she could reasonably accept laws against abortion (under another interpretation).¹⁴ Second, an interpretation through a might-counterfactual is not too tolerant, in making the reasonable acceptance of p too easy. It is not enough if, on some level of counter-factual abstraction, A would accept p .¹⁵

One of the tasks of justificatory liberalism is to describe the counter-factual situation S .¹⁶ But these details are not relevant here, as we already know that S has to comply with the internalism constraint. Let me instead turn to a different question, the various voluntaristic attitudes that we could have towards a law or institutional arrangement.

We should remember that, when speaking of “accepting p ”, p are not (simply) propositions, but (coercive) laws, institutions, and authorities. There is at least one major ambiguity when we speak of accepting a law or institution. First, we could mean that we *endorse* or *affirm* it. This is a strong notion, which will often entail advocacy of these views. It is in this sense that egalitarians “accept” egalitarian policies. However, there is a weaker notion, where accepting a law or institution means *tolerating* it, *acceding* to it, or *acquiescing* in it. In this sense, egalitarians might “accept” institutions that pursue center-right policies.

These two ways of speaking roughly correspond to accepting a law *as good* (or *just*), and accepting it *as legitimate*. In fact, this points us to an even greater variety of voluntaristic attitudes that individuals can express – such as accepting a law *as pragmatically useful*, *as imposing duties to obey* on us etc. It makes an obvious difference for (JL) whether reasonable acceptance of a law as good, or legitimate, or as possessing some other quality is required. This is an under-appreciated problem in the secondary literature. There is no such thing as reasonable acceptance *simpliciter*; rather, there is a set of various and very different voluntaristic attitudes.

“Accepting as legitimate” is likely to be the relevant attitude in the current context. The underlying idea is simple: if an individual accepts some coercive law *as legitimate*, then coercion based on that law has a prima facie claim to legitimacy, insofar as that individual is concerned. Requiring acceptance of the law as *good* or *just* – that is, requiring *affirmation* of the law – sets the bar

¹⁴ This is not possible if we interpret “could reasonably accept” through a would-counterfactual, such that A could reasonably accept that p iff, were A in S , A would accept p . Under a standard Lewisian interpretation, it is impossible that $(p \square \rightarrow q)$ and $(p \square \rightarrow \sim q)$. Many interpreters appear to favour such an analysis.

¹⁵ This is a problem if we interpret “could reasonably accept” through the possibility operator, such that A could reasonably accept that p iff there is some possible world where A accepts that p . Even if we play around with the sense of “possible”, this will much too often give the result that A could reasonably accept that p .

¹⁶ Note, however, that we have intuitions about what people could reasonably accept independent of specifying S .

unreasonably high, and also misses the point of legitimacy: legitimacy is about finding modes of living together, not of constructing the perfect form of government.

One worry with an accepting-as-legitimate standard is that individuals' assent might be already "tempered" with – assent to legitimacy might be due to pragmatic reasons, or a will to accommodate others. But this strategic character might diminish the normative significance of such acceptance. Second, there are circularity worries. Justificatory liberals accept a law as legitimate if *others* accept it as legitimate; but if these others are justificatory liberals themselves, this leads to an infinite regress. I will assume that these problems can be solved.

2.4.3 "Reasonably"

While the one major constraint on the interpretation of counter-factual situation *S* is the internalism constraint, the other is the notion that individuals are conceived as *reasonable*.

The word "reasonable" has several possible senses, but at least two major ones: a moral sense and an epistemic sense. In the *moral sense*, we assume that contractors are motivated by a minimal concern to propose "fair terms of cooperation".¹⁷ As a rough first pass, we can assume that citizens who are reasonable in this sense have some minimum concerns for morality, and are not entirely led by self-interest. It might also be that they need to fulfil some minimum requirements in moral reasoning – e.g., it might be that they are willing to universalise moral norms, and are willing to bear certain burdens.

In the epistemic sense, individuals are reasonable if they adhere to some minimum standards of rationality and inference, and if they actively seek out evidence, and respond to evidence in the right way. In an additional and related sense, we assume individuals to have essential relevant information about judging the situation, especially when it comes to factual claims. With regard to both epistemic and moral reasonableness, we impose some *minimum* threshold level of reasonableness, but we do not suppose that individuals are maximally or perfectly reasonable.

Assuming individuals to be morally reasonable will often violate the internalism constraint. The life the mafia boss would live if he were reasonable is quite different from his actual life, and unacceptable to him given his actual moral concerns. This might tempt us to say that we can ascribe reasonableness to individuals only insofar as that does not violate the internalism constraint.

That would be a mistake. The overall project of finding conditions for legitimacy is a *moral* project. This moral project should not be held hostage by the deeply immoral attitudes some citi-

¹⁷ An extensive analysis of moral motivation in contractualism can be found in Freeman 1991. I will return to the issues raised here in sec. 4.4.

zens possess. Thus, we are allowed to ascribe to citizens a minimum sense of moral reasonableness; a minimum concern for impersonal morality.¹⁸ While this violates the internalism constraint, we should not count this as problematic. (I will return to these issues in section 4.4.)

It will sound strange to some to say that the psychopath could reasonably accept his punishment, or that the mafia boss could reasonably accept anti-mafia legislation. This impression probably stems from the obvious breach of the internalism constraint that is implied in such ways of speaking. However, we should bite the bullet and accept such statements as meaningful and true claims. The awkward impression above claims leave might originate in mistaking “could *reasonably* accept” for “could accept”, or by mistaking “reasonable” for “rational” or “self-interestedly”. But reasonable acceptance is, despite the internalism constraint, a normatively coloured claim.

2.4.4 Rawlsian Constituency Restrictions

The current view differs from another common approach. Political liberals such as Rawls think that the *reasonable* acceptance of *reasonable* people is relevant for legitimacy. This approach has to define two notions: reasonable *people* and reasonable *beliefs*. It is based on a *constituency* restriction of (JL): only the dissent of certain people counts (see sec. 2.2). Both features are problematic, and make Rawls’s approach inferior to the current approach.

First, Rawls’s position produces confusion by introducing “reasonable” in two places (cf. Wenar 1995; Freeman 2007b, 345–6). On a natural understanding of these terms, unreasonable people can hold reasonable views, and reasonable people unreasonable views. These hybrid categories bring many problems with them. Why should it matter by whom reasonable beliefs are held? We have strong intuitions that claims in politics should be considered on their own merits, independent of the identity of the speaker.

For example, even if we think that the mafia boss should be ignored in setting up the criminal system, he might still hold important and right opinions on many other questions. But then we no longer need to distinguish between reasonable and unreasonable people, but only between reasonable and unreasonable belief. This collapses the current position.

Second, the Rawlsian constituency restriction means that there is an immediate need for a non-ideal theory which accompanies our ideal theory.¹⁹ Many citizens are not actually reasonable in the moral sense; we would need to say how they should be tak-

¹⁸ It is harder to decide whether the same is true for the epistemic sense of reasonableness.

¹⁹ See Friedman 2000; Kelly and McPherson 2001; Quong 2011, ch. 10.

en into account nevertheless, as it would be clearly repugnant to not extend the same rights and liberties to them. But this makes our theorising cumbersome, as it splits our theory in two halves (ideal and non-ideal).

Third, the strong binary separation of people into the “reasonable” and the “unreasonable” is unclear and arbitrary. Both moral and epistemic reasonableness come in degrees. Furthermore, particular individuals are often locally morally reasonable or unreasonable. The mafia boss is locally reasonable when it comes to family virtues. Thus, the reasonable/unreasonable split, which corresponds to the ideal/non-ideal split, is also misleading in that it paints a wrong picture of reasonableness.

Lastly, limiting dissent to the reasonable is strongly morally objectionable. We accept that “philosophical enlightenment should not be a precondition of moral status in a political society” (Kelly and McPherson 2001, 55). While the dissent of the unreasonable becomes morally significant again in the non-ideal part of Rawlsian theory, the exclusion of the unreasonable is highly problematic from a symbolic perspective.

2.4.5 Summary

I have now outlined the internalism and reasonableness constraints. This describes a family of positions: there is still much leeway about how one specifies what people could reasonably accept.²⁰ A major debate, for example, will be whether one should give the notion a more instrumentalist, Hobbesian ring – where what people could reasonably accept heavily depends on their desires and preferences; or a more reason-oriented, Kantian ring, where individuals are primarily led by their moral beliefs and concerns.

My critique of justificatory liberalism, however, is ecumenical enough that it does not rely on such details. The internalism and reasonableness constraints are enough to formulate the Basic Problem.

2.5 Acceptance and Rejection

2.5.1 The Standard Position

I formulated justificatory liberalism in terms of reasonable *acceptance*. But Parfit, Scanlon and Nagel prefer the notion of (the absence of) reasonable *rejection*.²¹ Thus, these commentators and others²² agree that there is a difference between the two standards. The general idea is that “what people can reasonably

²⁰ D’Agostino 1996 gives an excellent overview of the different kind of theory that are available at this point.

²¹ Scanlon 1982, 111–2; 2011, 137–8; Nagel 1991, 36; Parfit 2011. Cf. Moore 1996, 173.

²² D’Agostino 1996; Feltham 2007; Copp 2011, 246; Quong 2011, 209–10.

accept” is a less demanding standard than “what people can not reasonably reject”, i.e., there are more moral principles that people could reasonably accept than ones they could not reasonably reject.

Nagel formulates the common argument for this asymmetry claim:

“The range of institutions that people could accept without being unreasonable is far too broad—if one is willing to regard substantial levels of voluntary self-sacrifice as not unreasonable.” (Nagel 1991, 36; cf. Scanlon 2011, 137–8)

Nagel’s argument builds on the possibility of *supererogation*. It is morally permissible to voluntarily bear great burdens, or fore-go individual benefits, if this leads to some collective moral gain. Thus, we want to say that individuals could reasonably accept supererogatory acts. But this means that we would hold actual individuals to highly demanding standards if we use “could reasonably accept”: it would imply that we could sacrifice you because it would not be unreasonable for a highly supererogatory version of yourself to sacrifice herself. The “could not reasonably reject” phrasing is supposed to guard against such counter-intuitive implications, as there is one reasonable, counter-factual self which *does* reject that I am sacrificed.

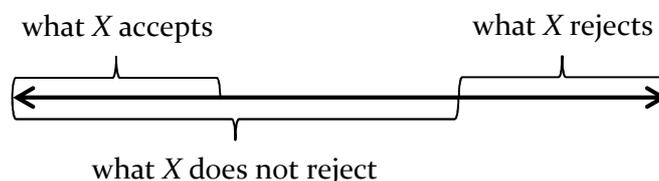
This strikes me as an unconvincing technicality, at least in the current context. First, there is an important difference between laws that *permit voluntary* self-sacrifice and laws that *mandate coercively enforced* sacrifice. Even though a counter-factual self might accept supererogatory acts, the later kind of laws might not be reasonably acceptable, even to her.

Second, what citizens do accept is often relevant to what they could reasonably accept (see sec. 2.3). Thus, if I know that I *do not* accept sacrificing myself, we will often say that I *could not* reasonably accept sacrificing myself. Third, the dilemma arises in its full strength only when we read “could reasonably accept” in a non-internalist sense which allows heavy idealisation, which I rejected.

2.5.2 Two Models

Still, it is worthwhile to consider models which explain possible asymmetries between acceptance and rejection. Let me sketch two.

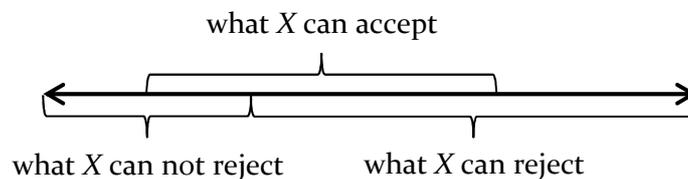
Model A. A first and natural model is highlighted in the following picture (where *X* is some individual):



Assume that there is a set of propositions (alternatively: laws, policies, etc.). We assume that an individual accepts and rejects some of them, and that, if someone accepts a proposition, she does not reject it. Without further assumptions, this implies that it is possible that there are some propositions that someone does not accept, but which she does not reject either. This is shown in the diagram above by the middle gap.

In this model, a justificatory liberalism that uses the “could (reasonably) accept” standard is necessarily stricter than a justificatory liberalism that uses “could not (reasonably) reject”, as there are fewer policies and laws that fall into the former class.

Model B. There is an alternative way to think about the relationship of acceptance and rejection, as can be seen in the following diagram:



The reasoning for this alternate picture might go as follows. *X* can reject *p* when there is any reason that speaks against *p*. (Alternatively, “when there is any *good* or *strong* reason that speaks against *p*”.) Thus, that *X* can *not* reject *p* implies that there is no (strong) reason that speaks against *p*. Symmetrically, *X* can accept *p* when there is some (strong) reason that speaks in favour of *p*. Note that this model, unlike model A, allows us to say that there are certain propositions (laws, ...) which an individual can accept, and at the same time can reject.

If we plug this model into justificatory liberalism, then there is no *a priori* reason to suppose that the “could reasonably accept” standard is more or less demanding than a “could not reasonably reject” formulation.

2.5.3 Two Conclusions

These models can be used to draw two conclusions.

First, the difference between models A and B allows us to further refine the different voluntaristic attitudes that can play a role in justificatory liberalism. With Pogge, we can say that model A understands “reject” to mean “dismiss” (2001, 124–6). It makes no sense to accept and dismiss a position at the same time. However, “reject” in model B means merely “disfavour” (Pogge). It is possible to favour and disfavour the same position at the same time, with respect to different features that it possesses.

Relatedly, “accept” in model A means “*conclusively affirm*”, whereas it merely means “*having (strong) reasons in favour*” in model B. (This difference in voluntaristic attitudes cuts across the differences I outlined in sec. 2.4.2.)

This leads to two versions of justificatory liberalism. On the first, “could reasonably accept” means “could reasonably conclusively affirm”; on the second, “could reasonably have (strong) reasons in favour”. Which formulation should we accept? The second model strikes me as too tolerant. To consider an analogy with actual consent, it is not enough to justify coercion to you that you see *some* beneficial aspects in it. To be normatively significant, you must agree that the beneficial aspects of my coercion outweigh its detrimental features. The same will be true for “could reasonably agree” standards.²³

Second, assuming that we accept model A, should we formulate justificatory liberalism in terms of acceptance or in terms of absence of rejection? The ultimate question is whether the absence of rejection has normative significance. It is not clear why the mere fact that some individual does not reject a policy is enough. After all, the absence of rejection might simply indicate indifference or undecidedness.

In criminal law, the absence of acceptance is often grouped together with rejection. Thus, as a crude approximation, an act would be rape *unless* it happens with the explicit or implicit agreement of the sexual partner. Similarly, contracts and promises usually become binding only through *actively* accepting them, not by not resisting them.

In law and morality, then, we regard as normatively significant that individuals express some *positive* voluntaristic attitude. Theorists who prefer “absence of rejection” standards often overlook this fact, or are falsely pre-occupied with technicalities about supererogation. However, a formulation in terms of acceptance fits much better with our considered convictions from other areas of law and moral philosophy.

2.6 Refinements

With the results from the last sections in mind, we can rewrite (JL) as

(JL₁) If some coercive law L is legitimate, then it is true for each citizen A that A could reasonably accept L (“could reasonably conclusively affirm L as legitimate”).

This formulation needs to be clarified in two respects: (1) the holistic nature of acceptance, and (2) the consensus—convergence distinction.

²³ We might strengthen model B such that “accept” means “have *very strong* or *conclusive* reasons in favour of”. But then it will only be very rarely the case that I could reasonably accept p , while I could also reasonably reject p , and model B will mostly behave like model A.

2.6.1 Holistic Acceptance

The focus of (JL₁) is a *particular* law. But there is an annoying feature about reasonable acceptance, its *holistic* nature. By this I mean that the reasonable acceptability of some law *L* will often depend on what other laws are in place together with *L*. For example, a law which regulates taxation might be reasonably acceptable to me if I know that tax revenues are used for infrastructure projects; but it might not be acceptable if I know that revenues are used to finance an unjust war.

Furthermore, if we apply (JL₁) on a law-by-law basis, we quickly run into aggregation paradoxes. For example, I could reasonably accept a law lowering taxes, and I could reasonably accept a law increasing social welfare spending; but I could not reasonably accept a *set of laws* that lowers taxes and increases spending at the same time.²⁴

To remedy these issues, we can turn to

(JL₂) If some *system of law S* is legitimate, then it is true for each citizen *A* that *A* could reasonably accept *S*.

But we might worry that (JL₂) makes legitimacy too easy to achieve. After all, it no longer demands the intense scrutiny which is directed towards particular laws under (JL₁). An intermediate position is

(JL₃) If some coercive law *L* is legitimate as part of system of law *S*, then it is true for each citizen *A* that *A* could reasonably accept *L* as part of *S*.

This formulation is supposed to avoid both the aggregation pit-fall, and the charge that legitimacy comes too easy. Because the “as part of system of law *S*” formulation is quite cumbersome, and usually not relevant to my discussion, I will usually stick to the (JL₁) formulation in this thesis.

2.6.2 Convergence—Consensus

(JL₁) leaves open the possibility that the same law is justified to different citizens for different reasons. But some commentators think that public justification requires something stronger: a *consensus* of reasons, rather than a mere *convergence*.²⁵ Put formally,

(JL₄) If some coercive law *L* is legitimate, then there is some reason or argument *R*, for which it is true

²⁴ This is an *intra*-personal aggregation paradox, and thus different from the *inter*-personal judgment aggregation paradox described by social choice theorists.

²⁵ The difference between convergence and consensus conceptions of public justification has been first highlighted by D’Agostino 1996, and further developed in Gaus and Vallier 2009; Quong 2011, ch. 9; Vallier 2011.

that each citizen could reasonably accept *L* on the basis of *R*.

The difference between (JL₁) and (JL₄) is often overlooked. (JL₁) presents public justification as a *conclusion- or outcome-based* ideal: *what* we agree on is public, but not *why* we agree on it. Here, public justification is not required to gauge the depths of reason. In (JL₄), however, this becomes important as well; here we adhere to a *premise- or input-based* ideal of public justification.

The guiding idea behind justificatory liberalism often is a concern for stability “for the right reasons”, and ideals of social and civic unity. The society necessary for legitimacy in (JL₁) might look dangerously like the “modus vivendi” that liberals like Rawls want to avoid. This has led some justificatory liberals to explicitly argue that (JL₄) in some form is preferable to (JL₁).²⁶

It is not clear why all individuals need to agree on laws for the same reason. The default is that policies could be justifiable for different reasons; respect for persons does not obviously require (JL₄). It is thus likely that the burden of proof lies with defenders of (JL₄). There is also a second reason to focus on (JL₁) in this thesis: (JL₁) is the more tolerant principle, and more policies will turn out to be legitimate on (JL₁) than on (JL₄). Thus, if we can argue against (JL₁), we will have an argument against (JL₄) as well.

2.7 Connections to Other Debates

With (JL₁) – and its more exact formulation (JL₃) – we now have a precise specification of (JL). Before I argue against it, let me note some connections between (JL) and two other debates.

2.7.1 Duty of Civility

A first connection is between (JL) and what Rawls has called the *duty of civility*. The principle says roughly:

Duty of Civility (DC). People in group *G* – e.g., legislators – have a duty to only appeal to public reasons when advocating, supporting, and deciding on coercive laws.²⁷

(A public reason is any principle, belief or argument that is publicly justified.) It might seem that (DC) is just the flipside of (JL). But note first the difference between (JL₁) and (JL₄). On the convergence approach of (JL₁), lawmakers do *not* need to appeal to shared (i.e., public) reasons.

²⁶ Brower 1994, 5; Quong 2011, ch. 9. For a defense of convergence, see Gaus and Vallier 2009; Vallier 2011.

²⁷ For a clear recent discussion, see Nussbaum 2011. Rawls’ most extensive treatment of (DC) is in Rawls 1999a. See also Solum 1993.

Second, even (JL4) would require only a very weak version of (DC). Only *lawmakers, in the act of law-making*, would be required to appeal to public reasons. In other words, group G would be very narrow, and the number of situations in which the duty applies tiny.²⁸ But this is hardly the relatively extensive “ideal of public reason” that many political liberals advocate, and the resulting duty would not be a very interesting one.

Third, there is an important difference between the acceptability of a *law* and the acceptability of *the reasons given for a law*. Imagine a group of moderate theocrats who install liberal policies of social justice, but for entirely theological reasons. While the moderate theocrats fail to justify their policies in terms acceptable to all, the policies *themselves* can be easily imagined to be publicly justifiable.

This difference is similar to the Kantian distinction between acting “out of duty” and “in accordance with duty”. The moderate liberal theocrats make law *in accordance with duty* – that is, their lawgiving stays inside the constraints of (JL) – but they do not act *out of duty* – that is, not on the basis of (DC).

This leads into a more general observation. Since Rawls, justificatory liberals have tended to overfocus on the reasons that are given for a law. But that we can accept the reasons given for a law is neither a necessary nor a sufficient condition for the acceptability of that law. Instead, we should accept *holism about justification*: whether a person could reasonably accept some law depends on the reasons given (by legislators, politicians, judges etc.) for it, but *also* on its content, practical implications, procedural pedigree, interpretation through the courts, and application through the executive.²⁹

For example, a law which guarantees certain human rights can be reasonably acceptable to citizens, *even if* the reasons given for it are not; inversely, even if the reasons given for an anti-terrorism law might be acceptable to all citizens, the fact that that law is likely to be abused for illegitimate purposes can make it the case that the anti-terrorism law can not be reasonably accepted. This is not to deny that sometimes an otherwise unproblematic law can become unacceptable to a citizen *because* it is based on reasons she can not accept. But it is unlikely and counter-intuitive that this will be universally so.

(DC) might express some worthwhile moral ideal such as an idea of civic community or civic friendship (Lister 2008), or might contribute to instrumental goods such as stability.³⁰ However, these are additional and independent concerns. The

²⁸ Some argue that the difference between law-makers and citizens should not be over-estimated (Waldron 1993). Even if this argument succeeds, it shows that we need an *additional* argument.

²⁹ This critique follows van der Burg and Brom’s (1999) discussion of neutrality, also adopted in Grotefeld 2006.

³⁰ Cf. Caney 1991 for consequentialist defenses of (JN).

Rawlsian emphasis on the reasons given for a law has distorted the debate about justificatory liberalism. Concern for public reason and a duty of civility are *not* core features of justificatory liberalism.

2.7.2 Justificatory Neutrality

There is another connected debate about whether a liberal state should be a *neutral* state, in particular between different conceptions of the good. Neutrality is often understood in terms of

Justificatory Neutrality (JN). A government ought not to appeal to any conception of the good, or any metaphysical or religious doctrine, in justifying public policy.³¹

(JN) differs from (JL) in several respects. First, (JN) is phrased in terms of what the state ought to do. It is not always clear whether advocates of justificatory neutrality want to make a claim about legitimacy or about some other normative property. Second, (JN) singles out “conceptions of the good”, a term which does not appear in (JL).

Nevertheless, the concerns underlying (JN) are similar to those that guide many liberals in accepting (JL), and we find a substantial overlap between advocates of (JL) and (JN). A reason for this overlap is that many liberals accept

(CG) (a) Conceptions of the good, as well as metaphysical and religious doctrines, are *inherently controversial*: (b) any law based on such principles is not publicly justifiable.

In other words, (JN) focusses on conceptions of the good because they are held to be *paradigm cases of the publicly non-justifiable*. If we add (CG) to (JL), it is natural to accept something like (JN) (cf. Caney 1998, 21). However, there are three reasons to treat (JN) as independent from (JL).

First, part (a) of (CG) is false, as can be seen in forms of moderate perfectionism (cf. Clarke 1999; Chan 2000). For example, claims about the advantages of physical exercise, the disadvantages of being addicted to heavy drugs, or the joys that family life can provide are not controversial, or at least they are not more controversial than claims in other areas.

Second, the idea of “conceptions of the good” is too vague to be of much help. While philosophers in the anti-perfectionist vein often assume that we have some good idea of what that idea means, this is far from clear. If “conceptions of the good” are simply meant to capture principles that are not publicly justifiable, then (JN) collapses into (JL) (cf. Klosko 2003). (Indeed, this

³¹ For formulations and debate of (JN), see Arneson 2003; Klosko 2003; Gaus 2003; Lecce 2008.

provides us with a helpful “translative technique” to appropriate many debates surrounding (JN).)

Third, (JN) once again suffers from over-focussing on the reasons given for a law. A government might act *on* perfectionist reasons, but that does not necessarily make the resulting law *itself* unacceptable. (Therefore, the inference from part (a) of (CG) to part (b) is false.)

Thus, if (JN) turned out to be false, (JL) could still be correct, and vice versa. Anti-perfectionism as expressed in (JN) is not among the core commitments of justificatory liberalism, and its plausibility stands and falls independently from it.

3 The Basic Problem

Justificatory liberalism is incompatible with two other liberal commitments. This is the “Basic Problem” which I will now describe.

3.1 Reasonable Disagreement

It is an empirical fact that individuals disagree about a multitude of issues. There is disagreement, for example, about

- how particular issues of policy should be decided,
- what justice requires generally,
- which forms of government and which laws are legitimate,
- what metaphysical and scientific theories best describe the nature of the universe,
- what makes a life meaningful and go well,
- which (if any) religious beliefs are true,
- how we should deal with disagreement in politics, and
- what standards of reason and evidence we should accept, both in the practical and theoretical realm.

Not only is there disagreement about these issues, but we commonly believe that these disagreements are *reasonable*. This can be seen especially clearly in the example of philosophy itself (or the humanities and social sciences in general). The variety of positions we encounter in political philosophy is breath-taking; any consensus on non-trivial issues in the near future is unlikely. Discourse in the philosophical community is highly reasonable; there is little distortion caused by money, power, or other political pressures; and strong norms are in place which favour pursuit of rational argument.

Disagreements in philosophy mirror disagreements among citizens. We can give a general explanation of why we encounter the latter.³² What it is reasonable for *A* to accept in the internal sense centrally depends on *A*'s past life experiences, including his education and family background, what he accepts as authoritative sources of reasoning and good standards of argument, which people he interacts with, and to which ethnic, social and religious group he belongs to. We can call this the “varying experiential situation of different people” (Rescher 1993, 3).

To this account we can add Rawls' “burdens of judgment”: the evidence that bears on laws and political decisions is complex, difficult, often incommensurable, and pluralistic (Rawls 2005, 56–7). Individuals have limited time and capacities to consider all evidence, or to form complete and fully coherent political

³² Other good accounts are given by Mason 1993; Gaus 1996, 17–108; Beson 2005, 17–120. See also Price 2000; Rescher 1993, chs. 4, 7.

views. As the amount of idealisation of individual reasoning we are allowed to ascribe to individuals has to adhere to the internalist constraint, we can not abstract these burdens away.

We can summarise this as

Reasonable Disagreement (RD). For any non-trivial law L (or: decision, principle, value, reason), reasonable disagreement among citizens in modern societies about L is to be strongly expected. (If there is reasonable disagreement about L , then there is one citizen who could reasonably accept L , and one who could not.)

I wrote “non-trivial”, as some trivial claims in morality and politics might not be contentious – possibly, because they are biologically hard-wired, or deeply ingrained in the evolutionary common core of social morality. Examples would be “pain is bad”, “murderers ought to be punished”, and maybe “family bonds deserve special protection” and “incest is wrong”. Some claims are also not subject to reasonable disagreement because no individual with a minimal amount of epistemic and moral reasonableness could reject them – for example, the modus ponens inference rule, or basic scientific facts.

Decidedly *non*-trivial claims would be “hate speech ought to be protected”, “a social minimum ought to be guaranteed to everyone”, or “everyone ought to have a right to organise in a union”. I will not attempt to make the distinction between the two sets of laws more determinate here; but enough laws are left in the non-trivial set that an attack on justificatory liberalism is possible.

I also wrote that disagreement is to be “strongly expected”. The sociological considerations I sketched do not work with deductive certainty for any law and value. Indeed, some philosophers have claimed to have found wide-spread agreement in actual societies. I return to such claims in section 4.1.

The reasoning I gave for the existence of reasonable disagreement rests on uncontroversial and general facts about modern societies. Since Rawls, (RD) is often labelled “reasonable pluralism”. But this label is unfortunate because (RD) makes no statement about the plurality of value, and is open to monistic interpretations of value (cf. Larmore 1994). Furthermore, (RD) does not entail value scepticism, or anti-realism in meta-ethics. (RD) is entirely concerned with what people *could reasonably accept*, constrained by *internalism*. This is compatible with saying that there is an objective order of values, and that perfectly reasonable deliberators would converge in their beliefs.

3.2 Anti-Anarchism

The third liberal commitment can be summarised as

Anti-Anarchism (AA). Many non-trivial laws are legitimate.

Most liberals are Liberals: they strongly believe that some form of social redistribution is a requirement of social justice, and that the liberal state can legitimately engage in such forms of redistribution. This can be seen in Rawls' work: while *Political Liberalism* advocates a form of (JL), Rawls unquestionably thinks that the distributive scheme outlined in *A Theory of Justice* is legitimate.

The argument for (AA) is not limited to social redistribution, or even justice. Any set of substantive political commitments that liberals have and which they regard as legitimate aims of the state imply (AA). In that sense, (AA) is merely a placeholder. While few liberals *explicitly* advocate or consider (AA), they will almost certainly have a set of more particular convictions about which non-trivial policies are legitimate that *entails* and *commits them* to (AA). Thus, it does not matter here which particular policies are thought to be legitimate, or how wide that set is.

With (JL), (RD) and (AA) in place, it is easy to state the *Basic Problem*: According to (JL), legitimacy requires public justification. But due to (RD), we can expect reasonable disagreement for all non-trivial policy issues. Thus, it is to be strongly expected that the majority, if not all, of proposed non-trivial laws will fail to be legitimate. But this contradicts (AA): Many non-trivial laws and policies, including the redistributionist policies required by liberal theories of justice, are held to be legitimate by liberals.

This problem is *basic* in three senses: First, it expresses a problem that has always intrigued liberals: explaining legitimacy under conditions of dissent. Second, the problem is not based on a conflict of any *particular* values, but on a conflict between two very high-level *sets* of values: a concern for legitimacy, and a concern for the pursuit of justice. Third and relatedly, this problem is more general and abstract than many other critiques that have been levelled against justificatory liberalism – e.g., that justificatory liberalism is unfair to religious believers, or fails to harness the epistemic advantages of diversity.

3.3 The Debate So Far

In *Political Liberalism*, Rawls explicitly acknowledged versions of (JL), (RD) and (AA). But Rawls wrote as if agreement on policies was achievable and unproblematic, and as if there were enough “public reasons” available to support a wide set of laws as legitimate. Rawls' critics in the 1990s, however, objected that Rawls had underestimated the problem (Waldron 1994; Sandel 1994). The charge became common that Rawls, despite his professed aims, had not taken serious (RD) (Caney 1998; Raz 1998); some remarks of the late Rawls did little to alleviate this impression (Rawls 2001). While Rawls and Rawlsians such as Larmore

and Nagel appeared to think it *obvious* that (RD) – and with it, the Basic Problem – could be overcome, some of their critics thought it *obvious* that political liberalism failed for the very same reason.

While insightful, this debate failed to fully come to grips with the Basic Problem. First, some of the early critiques knocked down straw men. Joseph Raz, for example, objects to (JL) that it would imply that the state has no authority to stop the “fanatical murderer”, as he would object to the authority of the state (1998, 40). This is an extremely uncharitable reading of (JL). It ignores the possibility that the murderer could *reasonably* accept those laws, and that laws against murder are one of the trivial laws that could be justified to everyone.

Another popular set of arguments observed that justificatory liberalism is itself controversial, and concluded that it is thus “self-defeating” (esp. Westmoreland 1999; Wall 2002). As I will argue below, this argument also misses the mark. Furthermore, discussions of (JL) were often been mixed together with discussions of anti-perfectionism and duties of civility – which, as I argued above, are separate issues.

Second, the majority of early responses focussed on Rawls’s political liberalism. For example, Waldron and Sandel objected that Rawls’s view implies that there can be no reasonable disagreement about justice. This might be correct as a critique of Rawls, but it is also correct *only* as a critique of Rawls. A preoccupation with the Rawlsian framework is still sometimes found in the secondary literature (e.g., May 2009). But we ought to tackle (JL) on its own merits. This is why I prefer the name “justificatory liberalism” – political liberalism is only one *species* of the *genus* justificatory liberalism.

Third and most importantly, there has been a renewal of justificatory liberalism in the 2000s, especially through the work of Gerald Gaus, David Estlund, and Jonathan Quong. These post-Rawlsian justificatory liberals have taken the critique and failure of Rawls seriously, and their frameworks have moved beyond Rawls in important ways.

Thus, the Basic Problem is still open, and should be raised anew. In doing so, we should avoid straw men and narrowing our debate to Rawls; and we should consider recent accounts of justificatory liberalisms in their strongest form. Aside from chapter 2, this will be achieved in the next chapter.

3.4 Reflexivity

Before turning to the Basic Problem, we can consider two arguments against justificatory liberalism which bear similarities to the Basic Problem, but which I reject.

Some justificatory liberals like Gaus advocate *non-reflexive* justificatory liberalism (Gaus 1996, 175–8; 2011, 226–8), as opposed

to *reflexive* justificatory liberalism. The former position denies, while the latter accepts

Reflexivity.³³ (JL) must itself be publicly justifiable, i.e., it must be the case that every citizen could reasonably accept (JL).

John Rawls' difficult remark that political liberalism applies the "principle of toleration to philosophy itself" (2005, 154) can be read to support reflexivity. Explicit support for reflexivity can be found in David Estlund's work.³⁴ Estlund defends the reflexivity of his "Principle of Qualified Acceptance",

(PQA) "No doctrine is admissible as a premise in any stage of political justification unless it is acceptable from all qualified points of view [...]" (2008, 53)

The word "doctrine" in (PQA) is naturally read as having universal scope, such that (PQA) applies to itself. Note that (JL) by contrast is formulated as a requirement on coercive laws. But (JL) is not a coercive law (nor even a law), and so we can argue that it falls outside the scope of (JL). This would be too quick, however; it is one of the guiding ideas of justificatory liberalism that *all political principles* ought to be publicly justifiable. Thus, there are internal pressures that (JL) be publicly justifiable, even if this does not follow from (JL) itself.

Many philosophers have been tempted by the following counter-argument to (JL):³⁵

(R1) Justificatory liberalism must be reflexive.

(R2) There is reasonable disagreement about (JL) – there is at least one citizen who could not reasonably accept (JL).

(R3) Thus, (JL) is "self-defeating".

(R4) Self-defeating doctrines ought to be rejected.

(R5) Thus, (JL) ought to be rejected.

This argument is short and simple, and attacks justificatory liberalism on an even more basic level than the Basic Problem. However, this argument is one of the straw men arguments I warned against. I first clarify a mistake about (R3), and then argue against (R1). I accept (R2).³⁶

³³ The label "reflexivity" has first been used by D'Agostino (1996, 61).

³⁴ Solum 1993, 735 also accepts reflexivity, and D'Agostino (1996) lists it as one of the desiderata for public reason.

³⁵ Cf. Raz 1998, 28–30; Westmoreland 1999; Wall 2002; Enoch 2009; Copp 2011.

³⁶ Estlund (2008, 61) and probably Rawls (2005, 61) deny (R2), by claiming that being a Rawlsian political liberal is a necessary condition for being (morally) reasonable (Enoch 2009) – an implication I find highly implausible. See also Moore 1996, 171; Lister 2010a; Copp 2011, 251.

3.4.1 Self-defeatingness

(R₃) draws the conclusion that (JL) is “self-defeating”, and this is language some critics have preferred (Wall 2002; cf. Raz 1998, 28–30). “Self-defeating” connotes that a doctrine is shown to be *false*. But the argument is not strong enough to show this. We have to distinguish *self-defeatingness* from *self-effacingness*.

To see the difference, we can turn to moral theory. The act-consequentialist imperative to maximise happiness is self-effacing (many argue) in the sense that individuals that act on it might fail to maximise happiness (cf. Parfit 1984, ch. 1). But this does *not* show that act-consequentialism, *as an objective criterion of what we ought to do*, is *wrong*; it merely shows that individuals *ought not to act on it*.

Similarly, the above argument does not show that (JL) is false. It only shows that, if we take (JL) serious, we ought not to act on it. (JL) can still be the correct principle of legitimacy; we just can not use it in public justification. Thus, the inference to (R₃) is mistaken.³⁷

One might reply that the argument still succeeds if we substitute “self-effacing” for “self-defeating”. But I do not think that self-effacing doctrines as such ought to be rejected (that is, (R₄) is wrong with regard to self-effacingness). There are analogous paradoxes in other areas of political philosophy – if we strongly believe in tolerance, we have to tolerate the intolerant, and if we think that neutrality is valuable we must be neutral even against the value of neutrality. But neither of these two contradictions should lead us to give up the value of tolerance or neutrality.

In a similar vein, if we think that valuable moral ideals support (JL), and that it applies to itself, then we might consistently accept the result that (JL) can not be used in justifying a law. By *not* using to it, we express our commitment to the deep value standing behind the principle – e.g., respect for persons. But this does not lessen or undermine our commitment to respect. On the contrary, we just consistently realise that value.

3.4.2 Anarchism Again

Thus, I reject the counter-argument as it is stated above. But we can still salvage something from the argument. Keeping (R₁) and (R₂), we can argue

(R₃^{*}) Thus, (JL) is self-effacing.

Self-effacingness implies, in the current context,

(R₄^{*}) We can not use (JL) to justify laws as legitimate.

³⁷ A similar argument to the same conclusion is made by Copp (2011, 249–50).

If we were to argue that law *L* is legitimate because it fulfils (JL) (and other conditions), then we would invoke a principle that is reasonably rejected by some citizens. Similarly,

(R5*) There will be reasonable disagreement about other principles of legitimacy.

This is a plausible assumption. From (R5*) and (JL) we can analogously infer that we can not use these other principles to justify policies as legitimate. (Note that our argument is not based on the claim that these other principles of legitimacy are self-effacing.) Together with (R4*), this leads to

(R6*) Thus, there is no principle of legitimacy we can use to justify laws as legitimate.

This implies

(R7*) Thus, there are no legitimate laws.

In other words, the problem with reflexivity is not self-defeatingness or self-effacingness; the problem is again that it leads to *anarchism*, a conclusion which conflicts with (AA). Thus, the best reconstruction of this counter-argument is analogous to the Basic Problem. The difference between the two arguments is that the Basic Problem relies on the *general* fact of disagreement, (RD), while the current argument only relies on the *specific* fact of disagreement about principles of legitimacy ((R2) and (R5*)).

3.4.3 Against Reflexivity

But even this improved argument fails, because (R1) is mistaken. Justificatory liberals do not need to accept reflexivity.

One argument *for* (R1) is that exempting (JL) from its own scope is arbitrary, and an unexplained asymmetry. Estlund writes that

“it would be odd to think that the truth (in this case the truth about the nature of justification), merely as such, could be sufficient for legitimacy here even though it is not sufficient anywhere else in political justification. [...] If we can eventually appeal to the truth merely as the truth in this way, it is no longer clear why we couldn’t appeal to it at an earlier point: why isn’t the religious truth also simply politically authoritative even despite reasonable disagreement?” (1999, 825; cf. Raz 1998, 31)

A first attempt to defend non-reflexivity is to claim that (JL) is a second-order principle, which regulates which first-order claims are admissible as political options or reasons.³⁸ We can appeal to the truth of second-order claims, but not of first-order claims. But what principled reason do we have to exclude sec-

³⁸ This is roughly Gaus’ reply (2011, 226–8). Cf. Colburn 2010 in a different context.

ond-order claims from the scope of public justification? This is again making the merely formal point that nothing *logically* requires (JL) to be reflexive.

For a better reply, we have to turn to the values which stand behind (JL). Assume here that this value is respect for persons (see ch. 5). Consider a few ways in which non-reflexivity could be considered disrespectful. First, we might think that denying reflexivity treats opponents of justificatory liberalism disrespectful, because it does not treat their views with the same respect as those of justificatory liberals. Justificatory liberals are allowed to appeal to the truth of (JL), whereas those disagreeing can not appeal to the (perceived) truth of their alternative principles of legitimacy. This is unfair.

Note first, however, that only the dissenters' views about *legitimacy* are not treated equally. In all other respects, reflexivity does not bias the case in favour of those citizens who accept justificatory liberals. The substantive disagreements that non-justificatory liberals have with policies remain untouched by this unequal treatment. Thus, we might argue that reflexivity is not required by respect for persons.

Second, if we accepted reflexivity, we would give non-justificatory liberals a veto on the principle of public justification. In some sense, this is giving individuals a veto against *equal respect itself*, insofar as we think that public justification is required by equal respect. The racist could object that certain groups ought not to be taken into account to determine legitimacy; and we would have to accede to these demands insofar as they are reasonable. But surely, if we believe in equal respect, we ought not to give people that kind of veto. Thus, there is even a sense in which denying reflexivity is *more* in accord with equal respect than accepting it.

Thus, non-formal reasons to include and exclude certain principles from the scope of public justification are provided by a substantive argument from respect for persons. It is *on substantive moral grounds* that laws regulating abortion are included in that scope, while (JL) is not. The arguments I sketched might turn out to be insufficient; but they at least show that it is in principle possible to make an argument against reflexivity.

If determining the scope of (JL) is a moral issue, then justificatory liberalism is committed to the truth of some political values. However, many justificatory liberals, impressed by Rawlsian concerns, want to deny that their view is based on accepting the truth of any moral value or principle. This explains part of the attractiveness of Reflexivity. If we can appeal to no value at any point in justifying (JL), then it is natural to think that (JL) applies "all the way down". There have been intricate debates about whether liberalism ought or can be fundamentally moral-

ly neutral in such ways.³⁹ I doubt whether it even *can*, and thus I will focus in this thesis on forms of justificatory liberalism that are based on substantive moral arguments, as will be seen in chapter 5.

I do not doubt that justificatory liberalism *can be* reflexive. But it is wrong to say that it *must be*, as is shown by versions based on substantive moral argument. Thus, justificatory liberalism can not be knocked down by simply observing that there is disagreement about it; there is still work to do for the Basic Problem. In any case, even if justificatory liberalism would have to be reflexive, the resulting problem would be very similar to the Basic Problem.

3.5 Baselines

We can use these observations to set aside another argument some have pressed against justificatory liberalism.⁴⁰ Implicit in (JL) is the idea of a *baseline*:⁴¹ deviations from a non-coercive, non-state baseline have to be justified; in the absence of successful public justification, we fall back onto that baseline.

But as Gaus has observed relatively early, we could imagine baselines other than a non-coercive one (1996, 164–5). For example, there could be an *equality* baseline: all deviations from an equal distribution of goods have to be publicly justified. This would turn the tables; now, the burden of proof is usually against state *inaction*. The Basic Problem would also vanish. Indeed, if we just set the baseline to broadly liberal policies, the conflict between liberal justice and legitimacy disappears.

Note also that there is reasonable disagreement about baselines: disagreement about how a given baseline is to be interpreted (e.g., what is implied by non-coercion?) as well as what baseline we ought to accept. But again, why should disagreement about the baseline be excluded from the scope of public justification? If we include baselines in the scope of (JL), then there no longer is a “safety net” if no reasonable agreement can be reached. In such situations, *neither* state action *nor* state inaction will be permissible, as we would be unable to publicly justify either. The result would not only be anarchism – the impermissibility of state action –, but also moral paradox.

A reply to this objection must again stress that justificatory liberalism is based on a substantive moral value – for example, equal respect. On the basis of this value, we would need to argue (1) for the non-coercive baseline – or whatever baseline we think follows from it –, and (2) against the inclusion of that baseline in the scope of public justification.

³⁹ See Raz 1990; Estlund 1998; Cohen 2003; Quong 2011.

⁴⁰ Christiano 1998, 457; 2009, 31–2; Wall 2010a; Lister 2010b; Arneson 2011.

⁴¹ Christiano 1998; cf. Raz 1988, 121–2; Gaus 2011, 485–6.

The Basic Problem

I can not sketch here how an argument for (2) would work. One general worry is that the kinds of reasons we give in favour of any particular baseline are just first-order reasons we give in normal political discourse. For example, an argument for liberty as a baseline might be based on the claim that liberty is a *prima facie* moral value. But arguments about the value of liberty properly belong into the scope of (JL), it seems. Thus, if baselines are simply expressions of prima facie values, we have strong reason to suppose that baselines fall inside the scope of (JL).

However, I will assume that a plausible argument for (1) and (2) can be given. Again, nothing *in principle* makes such arguments impossible, and thus we should not automatically presume that justificatory liberalism runs into this problem. We should also note that baseline-less forms of justificatory liberalism are also subject to the Basic Problem, as it also gives the result that the pursuit of justice is usually illegitimate.

4 Failed Solutions of the Basic Problem

With justificatory liberalism and the Basic Problem described in enough detail, I will discuss and reject six recent attempts to solve it. My discussion proceeds systematically, by considering “pure” versions of each answer. Many justificatory liberals use a mix of these answers. For reasons of space, I can not discuss the intricacies of each of these authors’ positions. However, I am confident that hybrid answers fail for reasons that can be seen from my objections to the pure versions.

4.1 Actual Agreements

One particular answer to the problem of disagreement is pursued by Klosko in his *Democratic Procedures and Liberal Consensus* (Klosko 2000; cf. Klosko 1997). Klosko advocates a “method of convergence”⁴²: We look at the opinions of actual citizens, who are reasonable in a comparatively weak, epistemic sense. Some intolerant fundamentalist Christians, for example, still count as reasonable.

By drawing on empirical studies in opinion research, Klosko argues that (1) even such weakly reasonable citizens will agree on a core set of abstract rights and political principles – for example, freedom of speech; and that (2) they will agree that the exact interpretation of these abstract rights should be left to “democratic procedures”. (The second aspect touches on the issue of “meta-agreement”, which I will discuss later.)

Klosko is not alone in his optimism. Gaus similarly draws on empirical studies to show support for some abstract principles (e.g., 2011, 356–9). Much of Rawls’ political liberalism is premised on the availability of a historically formed “public political culture” from which support for abstract liberal principles can be drawn (Rawls 2005, 8–9, 13–5, 43). In the United States, for example, this public political culture would probably compromise the constitution and some central Supreme Court cases, as well as historically important books and speeches, such as the Federalist papers or the Gettysburg address.

While these observations are about actual consent, they are directly relevant to the question of what people could reasonably accept. Weak internalism implies that what *A* can reasonably accept depends on light idealisations of *A*’s current epistemic situation and moral motivation. Thus, what people actually accept possesses some relevance.

⁴² The label is not misleading, even though Klosko does not make the “convergence—consensus” distinction – Klosko does not require citizens to agree on principles for the same reason.

Despite some minor quibbles on how Klosko reads empirical evidence – which is much more inconclusive than Klosko thinks (Archard 2000; Weithman 2002) –, we should be charitable and acknowledge that there currently is a relatively wide consensus on relatively vague and abstract liberal principles in most Western democracies. This is due to a shared history and experience, a common cultural background, strong institutions and educational systems, and the influence of existing constitutions and laws. The actual shape of the agreement varies from country to country – Germany, for example, has an overwhelming majority that disfavors the permissibility of hate speech, while the opposite is the case in the United States.

In reply we should note first that great, even overwhelming majorities are not enough in justificatory liberalism – we need *everyone's* reasonable acceptance (cf. Eberle 2002, 210–1). Thus, we would need to show that even those who currently disagree with a system of laws could reasonably accept it. This is far from a simple job. After all, those who strongly disagree with mainstream opinions are often those that have especially strong and considered reasons why they hold a deviating position. Widespread agreement is likely to be more due to indifference and lack of critical engagement.

Second, remember that we are interested in non-trivial agreement. We should readily admit that citizens can agree on some very general principles, rights, and liberties – e.g., the right to a fair trial, or the desirability of punishing criminals. But our focus here is on much more controversial coercive laws – e.g., substantive material redistribution, affirmative action, or the reparation of historical injustices. Little agreement is to be expected on such non-trivial laws; but such laws which aim at social justice are often of central concern to liberals.

Third, even if these last two issues can be resolved, the current solution fails for a more basic reason. Liberalism should have answers even for regimes in which no institutional, historical or cultural background consensus exists on which liberal institutions can be built. Importantly, the pursuit of justice is legitimate under such circumstances as well. Even stronger, it is *especially* important that the state is morally permitted to support democratic institutions and liberal education if wide-spread support on these issues is currently lacking.

Even if *actual* agreements would make *actual* justice-oriented policies legitimate, liberalism as a position in political philosophy ought to have answers for all empirically likely circumstances, including worst-case scenarios. In the language of social choice, we can call this assumption “unrestricted domain” (e.g., List 2011). It is no accident that many liberals like Rawls focus on countries like the United States or European countries, who have settled agreements on many issues of policy. But there are many younger and more contested democracies, for exam-

ple in Egypt, Turkey, India, Ghana, and Russia, for which this is not true.

This reply spells doom for any related position that tries to answer the problem of reasonable disagreement on the basis of actual agreement. Alasdair MacIntyre, for example, thinks that the disagreements inherent in liberalism can be overcome through a substantial, shared “ethical tradition” (MacIntyre 1988). And many “deliberative democrats” hope that disagreements can be reduced over time through liberal education, civic engagement, and wide-spread public deliberation (e.g., Manin 1987; Habermas 1998).

These two positions agree in their hope that reasonable disagreement can be overcome through “social engineering”, either exogenous or endogenous to democracy (List 2011), in the long run. But even if we assume that these forms of social engineering *would* succeed in reducing disagreement, the legitimacy of policies that advocate social justice should not contingently depend on the availability of reasonable agreement. That an ethical tradition or liberalism might ultimately convince everyone is irrelevant to the *here and now*, where disagreement looms, but where policies of social justice are *just as* legitimate. Thus, the availability of such instruments does not alleviate the Basic Problem.

4.2 Procedural Meta-Agreement

Another strategy to resolve the Basic Problem is the possibility of meta-agreement.⁴³ We might “agree to disagree” – we might cast our first-order disagreements aside, and agree on a certain procedural solution on resolving our disagreements. For example, we can resolve contested issues via simple majority voting, or through a judicial system, or some more complicated institutional structure. Other “agree to disagree” solutions also merit mention, such as moral compromise and accommodation.⁴⁴

This argument can be coupled with other arguments. Klosko claims that there is large *actual* agreement that disagreements should be resolved through democratic procedures (2000, ch. 5). The late John Rawls also claims that we should “simply vote” among different reasonable political conceptions of justice (1993, liii–liv).

In support of the “procedural meta-agreement” solution, we should first observe that this is indeed our political reality. Debating all actual disagreements is usually time-consuming and inefficient, and so we accept various procedures to bypass such disagreements – delegation, voting, or even lot.

⁴³ List 2002; Schwartzman 2004; Dryzek and Niemeyer 2006; Niemeyer and Dryzek 2007.

⁴⁴ On compromise and public reason, see May 2005; Lister 2007.

However, observations from such every-day conflict resolution procedures can be misleading. In these questions we deal with pragmatic questions of comparatively little importance. Similar remarks can be made about compromise: compromise is easy when it comes at little cost. But when we make basic political decisions, we decide quite fundamental matters, which will determine our whole lives and identities.

In general, “agree to disagree” solutions only work when I have an exclusionary reason to accept the results of the conflict resolution procedure (Schwartzman 2004, 213–4). If the chosen second-order procedure selects *L* as a result, then *L* must be reasonably acceptable to me, no matter what first-order reasons I might have which speak against *L*. But the more weight the first-order decisions have, the less likely will it become that I have such an exclusionary second-order reason. Usually individuals can accept the outcomes of some procedure if it comes to city ordinances such as traffic laws; but when it comes to laws which regulate (say) abortion or conscription, many citizens could reasonably accept the outcomes of a second-order procedure only insofar as that outcome tracks their first-order reasons.

In other words, the plausibility of a second-order agreement rests on a form of pure proceduralism: procedural values crowd out all substantive values. But this is an unrealistic picture, which I already rejected with regard to legitimacy (sec 2.2). It will often be true that citizens have a *duty to obey* the outcomes of *actual* democratic procedures; but this leaves untouched the claim that it is often the case that they could not reasonably accept the outcomes of that procedure (more on this in sec. 4.5.3).

As a second point, we should note that there is no such thing as a “democratic procedure” simpliciter. For example, there are strong, reasonable disagreements about the ways in which legislative decision-making ought to be constrained by judicial review (Waldron 2006), whether democracy ought to be merely “aggregative” or whether it should be “deliberative”, and whether we should prefer the “Westminster system” of government to the European “consensus democracy” system (Lijphart 1999). There are also plausible non-democratic decision procedures such as lottery voting systems (Saunders 2010).

In other words, we should expect as much second-order disagreement as first-order disagreement. Even if it were true that we have a strong reason to accept all outcomes of some conflict resolution procedure, different individuals have reason to accept different procedures. To try to resolve *this* disagreement by some further higher-order process leads to a vicious regress.

The acceptability of procedures heavily depends on the acceptability of the outcomes it brings about. As an example, contractualists like Habermas or Scanlon are much more likely to fa-

your broadly consensual forms of democracy, as such a procedure tends to bring about results that achieve everyone's acceptance. On the other hand, perfectionists will prefer more elitist forms of government. Thus, our acceptance of second-order procedures depends on what we think good first-order outcomes are. We can expect first-order disagreements to "leak into" disputes about second-order procedures. But then this higher order does not provide any advantages over the lower orders.

For these reasons, procedural meta-agreement does not promise a significantly better solution to the problem of reasonable disagreement.

4.3 Asymmetries

A third answer to the problem of reasonable disagreement is to stake out a certain area in which reasonable disagreement can be avoided, or where it is not vicious. *Political* liberals have thought that the *political* escapes vicious reasonable disagreement. The liberalism of Rawls, Nagel and Barry focus on such asymmetries (Nagel 1987; Rawls 2005 [1993]; Barry 1995). Habermas' discourse ethics trades on a very similar distinction between the *moral* and the *ethical* (Habermas 1998; Forst 2007, ch. 3), where the former is supposed to be much less subject to reasonable disagreement than the latter.

These arguments have come under a barrage of criticism (Sandel 1998, 202–210; Chan 2000; Caney 1995; Clarke 1999; Waldron 1999; Lecce 2008, ch. 6). Various authors have found it implausible, on both empirical and a priori grounds, that there should be any such asymmetry. Even if one admits a gradual difference between the political and the non-political, controversies in political matters still are deep and stable enough to thwart any reasonable agreement.

While I ultimately agree with the critique of the authors cited, political liberals have advanced some promising arguments that support asymmetry. I will first turn to Rawls' support for asymmetry, and then to a recent, Rawls-inspired defence of asymmetry by Jonathan Quong.

4.3.1 Rawls

Rawls thinks that we aim for agreement on "political conceptions of justice". These have three characteristics: (1) they have limited *scope* – they only deal with matters that pertain to the "basic structure" of society; (2) they have limited *depth* – they leave religious, philosophical and metaphysical issues un-

touched;⁴⁵ and (3) they have limited *content* – they only draw on ideas contained in the “public political culture” (see above).

We should admit that features (1)-(3) indeed *reduce* the amount of reasonable disagreement that we can observe in political conceptions of justice. This is straightforward with regard to features (1) and (2): if we have to find agreement on less, and with lesser depth, we are more likely to agree. Feature (3) is again a sign of Rawlsian over-emphasis on the reasons given for laws (see sec. 2.7.1); it is not a priori clear whether laws become more acceptable if the reasons given for them are only public.

However, these features are unlikely to *eliminate* reasonable disagreement. It is a sociological and empirical fact that individuals disagree on a wide range of political issues, even on the basis of supposedly shared ideas and without invoking metaphysical or religious foundations. At the very least, the current position does little to explain these occurrences away.

Second, assume that individuals *can* agree on a particular set of *political* reasons. Thus, all individuals agree on what political reasons there are, how strong they are, and what they imply. Note how strong especially the last two requirements are. However, even if we grant this, this does *not* eliminate reasonable disagreement about which laws are acceptable. It might be, for example, that political reasons *P* support coercive law *L*, – and by stipulation, all citizens could reasonably accept *P* and the inference from *P* to *L*. But that does *not* exclude the possibility that individual citizens have further, non-political reasons *Q* that speak against *L*, such that they could not reasonably accept *L*.

For the Rawlsian inference to be correct, it has to be true that

Only the Political Counts (OPC). If *A* could reasonably accept law *L* merely on the basis of (shared) political reasons *P*, then *A* could reasonably accept *L*.

This claim would be true if it were the case that individuals only or primarily had political reasons. (OPC) would be trivially true if we defined “political reasons” as those reasons that are relevant to deciding political matters. But remember that we defined political reasons above as being limited in scope, depth and content, and thus we can not make that inference.

Rawls argues for (OPC) by claiming that *reasonable* citizens would only be concerned with political reasons. But on an epistemic sense of reasonable, (OPC) is wrong. In fact, citizens would be *unreasonable* in the epistemic sense if they did not take all reasons into account when deciding on which laws they ought to accept.

⁴⁵ Scope is different from depth. You can argue for specific, narrow political principles on the basis of deep metaphysical foundations, and for wide-ranging political requirements on the basis of shallow foundations.

If we turn to a moral sense, then (OPC) starts to look question-begging. *Why* would reasonable citizens *only* invoke political reasons? The presumptive answer is that it is a morally necessary requirement for legitimacy that policies are reasonably acceptable to everyone. But that means that we would have come full circle. Independent of such an argument, it is hard to imagine what could justify (OPC). Together with the other critiques I made, this shows Rawls' argument for asymmetry to be a failure.

4.3.2 Quong

Jonathan Quong has made a better argument for asymmetry in his recent book, *Liberalism without Perfection* (Quong 2005; 2011, ch. 7). Quong's argument can be reconstructed to proceed in three steps:

(1) Quong distinguishes two forms of disagreement, *foundational* and *justificatory* disagreement. If *A* and *B* have a *justificatory disagreement*, then *A* and *B* have some shared (normative) premises and principles, and they only disagree about which inferences follow from these principles. In foundational disagreement, on the other hand, *A* and *B* "disagree at the level of ultimate convictions or principles", that is, "about what the standard of justification [itself] should be" (Quong 2011, 205). Quong's distinction is similar to a suggestion by Michael Sandel, who distinguishes disagreements about the application of given principles, and "foundational" disagreements about the principles themselves (Sandel 1998, 204).

(2) Quong claims that there is an area where disagreements are always justificatory – the area of the political. According to Quong, there is a small set of political principles and ideas which every reasonable citizen ought to agree with. They are quite Rawlsian (Rawls 2005, 9, 15–22): the "fundamental organizing idea" of "society as a fair system of social cooperation", plus the two "companion ideas" of "citizens as free and equal persons" and of a "well-ordered society". That everybody could reasonably accept these principles is not an empirical claim, but due to a morally substantive reading of reasonableness. Thus, in all *reasonable* political disagreements, we will share at least the values of equality and liberty, for example.

(3) Quong argues that the exercise of power is unproblematic if it is in the face of justificatory disagreement. Thus, if *A* and *B* have a reasonable justificatory disagreement about law *L*, then coercion on the basis of *L* against both *A* and *B* is still legitimate. This is because Quong thinks that in a reasonable justificatory disagreement I could reasonably accept the position of my opponent: as he proceeds from the same premises, I could accept the conclusion of his argument (given that his argument adheres to some standards of argument). Thus, all claims that are based on reasonable interpretations of and arguments from the three basic political principles will be reasonably acceptable to me, as they will be to everyone.

This is a sophisticated argument for asymmetry. In my critique, I will focus on claims (1) and (3). Moralised readings of “reasonable” will be discussed in the next section, and I will leave aside that premise until then. Note that if Quong claimed that (2) is based on an empirical argument about what people actually accept, his current defence would fall for the kind of objection against empirical agreement that I outlined above.

Let me start with (1). First, Quong’s distinction between “justificatory” and “foundational” disagreement is overly vague, and much analytical nit-picking could be directed in this direction. For example, it is possible that a liberal and a non-liberal *can* agree on “ultimate convictions” and “standards of justification”, while *disagreeing* on the liberal commitments that Quong mentions. This is because liberalism hardly is an “ultimate conviction” – despite what some political philosophers think. Inversely, it is possible to agree on Quong’s liberal commitments while having different standards of justification or ultimate convictions.

More problematically, Quong’s distinction presupposes a foundationalist, hierarchical view of moral beliefs. This implies that the distinction could no longer be operationalized in coherentist views in moral epistemology. In a coherentist view, we can not simply say that individuals have “ultimate premises”, or that multiple individuals have “shared premises”: our moral beliefs are not structured in a “premise—conclusion” structure, even if moral arguments are. Instead, we hold various moral beliefs on different levels of generality and scope, none with any necessary priority over the others. Thus, even if I agree on the more abstract levels with someone, our disagreements about more particular cases can, and often will pry us apart.

While I can not defend it here, I believe that an adequate moral epistemology will be coherentist in many regards. It is a serious shortcoming of Quong’s approach that such an epistemology is not compatible with his explanation of disagreement.

Second, Quong’s justificatory agreements are mostly verbal, and as such quite powerless. Consider the classic opposition between Nozick and Rawls. Nozick *never* disagreed with equality and liberty as values, and would also have been able to accept versions of society as a system of cooperation and the idea of a well-ordered society. In this sense, Rawls and Nozick had shared foundational premises. But to say that the disagreement between Rawls and Nozick is justificatory for this reason is not very informative, and false.

True, both Rawls’s and Nozick’s positions find their shared basis on some Kantian platitudes. But that shows little, if anything. We can reformulate most disagreements in such verbal ways, and make them disappear. One can observe this, for example, in presidential commencement speeches. The lofty appeal to vague and ambiguous values in these speeches makes possible disa-

greement quite difficult. As further examples, the atheist and theist agree that “there is an external reality” and that “we should assess evidence rationally”; the virtue ethicist and the utilitarian can agree that “human welfare matters and should be the sole basis of ethics”. While this gives the veneer of agreement, the surface appearance is quite misleading.

On the basis of this, claim (3) is unlikely to be true. The fact that there is some abstract verbal basis which makes our disagreements appear justificatory does not make it the case that I could reasonably accept any opposing position that can be formulated on the basis of that same verbal basis. Merely because Nozick appeals to the same Kantian values and principles which are also found in Rawls, does not make his position reasonably acceptable to Rawls, or vice versa.

Third, Quong’s position also fails for a reason similar to the reason of the failure of (OPC). Assume that *A* and *B* have shared premises *P*, and *on the basis of these shared premises*, *A* could reasonably accept *L*, even if *A* disagrees with *L* on his own reading of *P*, and *L* only follows on *B*’s reading of *P*. But from this it does *not* follow that *A* could reasonably accept *L*. For this to be true, it would have to be true that

Only Shared Reasons Count (OSRC). If *A* could reasonably accept law *L* merely on the basis of reasons for which he share premises with others, then *A* could reasonably accept *L*.

But *A* can have reasons that he does not share with everyone, but which are relevant for deciding his reasonable acceptance. For example, *A* might have utilitarian, or theistic, or perfectionist reasons, which he will weigh against those “political” reasons which are only subject to justificatory disagreement. Put otherwise, even if it were true that there is a set of reasons about which disagreement is merely justificatory, the fact that there will be non-justificatory disagreement about *other* reasons undermines the hope that laws will be generally reasonably acceptable to everyone.

For these reasons, Quong’s defence of asymmetry fails. Quong presents the best and most sophisticated defence of asymmetry, and so it is likely that asymmetry in general fails as an answer to reasonable disagreement.

4.4 Moral Content

Another common way to avoid reasonable disagreement is to use a strongly moralised version of what individuals could *reasonably* accept. For example, we might claim that citizens who do not accept political principles that guarantee a minimum provision of goods to everyone are unreasonable; thus, their disagreement does not count.

Heavily moralised versions of reasonable acceptance are used by many philosophers. Scanlon assumes throughout in his contractualism that individuals are strongly motivated to seek principles that no one else could reject, and many authors have thought that the idea of the reasonable does the main work in Scanlon.⁴⁶ Nagel writes that “what it is reasonable to reject is a moral issue, ‘all the way down’” (1991, 39). Many authors advocate similar positions in the political realm.⁴⁷

One of the most prominent, again, is John Rawls. We can identify two and a half main elements of what it means for a person to be reasonable in Rawls: (1) the person is “ready to propose and willingly abide by principles and standards that are fair terms of cooperation”; (2) she recognizes the burdens of judgments, and (2b) she also accepts “[the] consequences [of the burdens of judgment] for the use of public reason in directing the legitimate exercise of political power in a constitutional regime” (Rawls 2005, 54).⁴⁸ By (2b) Rawls probably means that reasonable people will be justificatory liberals.

A simple counter-argument is to point out that strongly moralised readings of “reasonable” violate the internalism constraint of “could reasonably accept”. That answer would be too quick, however. As I argued above, we can ascribe moral reasonableness to individuals beyond the internalism constraint, because otherwise we give murderers, egotists, and uncooperative members of society a say.

Still, an objection along these lines succeeds. First, I will describe the form of reasonableness that we ought to ascribe to individuals as *minimally objective, motivational* reasonableness. I argue that on this form of reasonableness enough reasonable disagreement persists for the Basic Problem to arise. Third, I claim that ascribing stronger form of reasonableness to individuals no longer yields a plausible justificatory liberalism.

4.4.1 Forms of Moral Reasonableness

The reasonableness that Scanlon and Rawls describe is *motivational* reasonableness. Motivational reasonableness is a certain disposition to interact with others, to react to the concerns, interests and arguments of others, and to cooperate in certain ways. Reasonable individuals are sensitive to impartial moral concerns, and they are willing to impartially step back from their personal biases and self-interested perspective to find a solution to political problems. This motivational reasonableness is not described through any particular set of doctrines or propositions; it is more akin to a character virtue than a set of rules. This will become important later on.

⁴⁶ For debate, see Freeman 1991; 2006; Moore 1996.

⁴⁷ Wenar 1995; Lecce 2003; 2008; Brettschneider 2007.

⁴⁸ A full list would include further elements, which I omit here. Cf. Wenar 1995, 37; Lecce 2008, 214–222; Freeman 2007b, 345–351.

There is an ambiguity in Rawls's and Scanlon's notion that reasonable individuals seek "terms others could not reject", which also carries over into phrases that reasonable individuals seek "fairness" or "justice", or are "sensitive to moral concerns". First, we can interpret this subjectively (or "de dicto") to mean that individuals seek terms *they perceive as* fair, or as ones others could not reject. Second, this can objectively (or "de re") be read to mean that individuals seek terms which are *in fact* fair, or ones other could not reject.

Subjective moral reasonableness is not an empty notion – it is not definitionally true that everyone is reasonable in this sense. Some individuals are unconcerned with fairness or justice, and only propose terms of cooperation which suit their own interests. If we ascribe such reasonableness to some citizens, this will already be enough to violate the internalism constraint.

However, purely subjective reasonableness rules out very few opinions. Many Bolsheviks in the 1920s and 1930s thought, for example, that state terror was necessary and legitimate, and Lenin and Trotsky argued for this in their theoretical works. It is not unrealistic to think that the Bolsheviks genuinely thought these policies to be what legitimately ought to be done. Even if you reject this example, many clearly repugnant opinions are compatible with purely subjective moral reasonableness.

But in determining which policies are legitimate, we ought not to be held hostage by such morally perverse opinions. Some objective elements have to be mixed into what counts as reasonable. But how can we draw a principled line here, without moving to the opposite end, that is, accepting a purely objective interpretations of reasonable?

We should remember that we operate in a non-reflexive justificatory liberalism based on substantive moral foundations: equal respect for persons. This suggests the following

Reasonableness Constraint (RC). Beliefs or systems of beliefs that are fully incompatible with any plausible interpretation of equal respect for persons are unreasonable.

If we denied (RC), then we would claim that equal respect for persons is a valid and true moral principle of central importance, but also a principle that some citizens can reasonably reject. This would be an odd combination of claims. We are not epistemically neutral with regard to the values and principles that underlie (JL); there are internal pressures inside non-reflexive justificatory liberalism to accept (RC).⁴⁹

With (RC), we move beyond a purely subjective reading of moral reasonableness. (RC) rules out moral opinions that are morally perverted, vicious, or otherwise subject to strong moral error,

⁴⁹ This is a very liberal interpretation of an argument by Lecce (2008, 170).

even if these opinions are the outcome of a perfectly honest individual's attempt to impartially find the moral truth. Still, the weak formulation I chose ("*fully incompatible with any plausible interpretation*") leaves much leeway for individual variation. Thus, I call a motivational notion of reasonableness which is constrained by (RC) *minimally objective*.

4.4.2 The Persistence of Disagreement

Minimally objective reasonableness allows us to screen out political opinions that are voiced due to viciousness, partiality, ignorance, strong moral error, or egoism. Under this reading, no one could *reasonably* accept genocidal politics, arbitrary terror, systematic torture of the innocent, or a system of law that treats equals unequally – because none of these systems are compatible with equal respect for persons. Thus, there will be reasonable agreement on many trivial, and maybe even some non-trivial laws.

However, wide-ranging disagreement about many non-trivial laws and positions will persist. We can consider here an example, libertarianism. It is not hard to imagine a libertarian who has the correct subjective moral motivation – that is, who honestly thinks that libertarianism is an expression, and the only adequate expression, of equal respect. He can also think, without obvious moral error, bias or viciousness, that libertarianism presents a fair and just system of cooperation.

Does libertarianism violate (RC)? Is it morally pervert, or in contradiction to equal respect for persons? This would require a discussion of its own, but the *prima facie* answer is no. Libertarianism strongly respects the autonomous choices of individuals; it assigns strong rights and freedoms to them. If Nozick is right, libertarianism even acknowledges rights as absolute side-constraints: as absolute vetoes that citizens have against any social choice which would interfere with certain aspects of their lives. There is then at least one interpretation of equal respect with which libertarianism is compatible; it is not ruled out by (RC).

Similar arguments could be pursued with regard to egalitarian and utilitarian policies. Dworkin's egalitarianism is explicitly intended to be an interpretation of equal respect for persons. The case for utilitarianism is controversial, but some utilitarians have claimed that utilitarianism is compatible with equal respect for persons (Cummiskey 1990), or even that it respects persons better than non-consequentialist moral theories (Pettit 1989). Thus, egalitarianism and some forms of utilitarianism are compatible with some interpretations of respect for persons.

This should make it clear that enough reasonable disagreement remains under (RC) to state the Basic Problem again. Deep and wide disagreement remains, even after we have filtered out the opinions of Bolsheviks, fascists and theocrats.

4.4.3 Substantive Reasonableness

But why not use a notion of reasonableness which is stronger than this? Assume that in addition to (RC), we also accepted

(RC₂) It is *unreasonable not* to accept a guaranteed social welfare net for everyone.

This makes much more specific prescriptions than (RC), and rules out a much larger set of laws and opinions. Call a position in which we accept a number of reasonableness constraints as specific as (RC₂) *substantive reasonableness*. If we accept substantive reasonableness, the Basic Problem will no longer be a very great challenge. However, there are three reasons why substantive reasonableness should not be accepted by justificatory liberals.

First, we can press a *redundancy objection* against substantive reasonableness (cf. Quong 2005, 310–1; 2011, 201).⁵⁰ If our notion of reasonable acceptance is that heavily prefigured in a moral sense, then all work is done by figuring out what is reasonable. But justificatory liberalism in the first place was intended to take most of these decisions away from the hands of the political philosophers, and open them up to genuine moral argument in the general public. A justificatory liberalism which shies away from this is not necessarily wrong, but uninteresting.

Furthermore, if all the work is done by a substantive notion of reasonableness, the idea of individual *acceptance* – the voluntaristic element in “could reasonably accept” – in particular starts to look redundant. What use is there to consider what individuals would accept, if we operate with such strong conceptions of reason? In other words, our notion of reasonableness crowds out the internalism constraint. But then we can no longer plausibly claim that what individuals could reasonably accept has anything to do *with them*. We should repeat here that it is inevitable that the internalism and reasonableness constraints will clash; but in the current position, internalism is practically eliminated, which strongly speaks against this position.

More generally, there are respect-based reasons against substantive reasonableness. While we ought not to give citizens a veto against equal respect, we ought to allow them to develop their own interpretations of what respect for persons means and implies. Substantive reasonableness does not allow for such leeway, and thus clashes with respect for persons.

By contrast, motivational reasonableness does not crowd out the internalism constraint in such a way. Motivational reasonableness is “procedural”: it describes a way how individuals react to moral evidence, and interact with others, behave in moral

⁵⁰ This is related to redundancy objections levelled against Scanlon’s contractualism. See Suikkanen 2005 and Southwood 2010, ch. 3 for pointers to the literature.

argument, and live together with others in a society. But it leaves room for individuals to fill this in through a variety of individual commitments, beliefs, and identities.

Second, (RC₂) and claims which are similarly specific strike me as wrong on a natural reading of “reasonable”. “Reasonable” is unlike “true”; rough synonyms for it would be “plausible” or “viable”. For example, even if we think libertarianism to be false, it is much stronger to say that it is implausible, that is, that nothing can be said which *prima facie* supports it. One has to have a very clear and unconflicting set of moral intuitions to support this drastic claim.

The notion of a “moral person” is relevantly similar here. We might think that utilitarianism is a wrong moral theory; still, it is a mistake to call an enthusiastic follower of utilitarianism an “immoral” or “non-moral” person, as long as he does not commit grave moral errors.

Third, what reason do we have to accept (RC₂)? We can not accept (RC₂) on the basis of the claim that a social welfare net is legitimate, as this is the very question we are trying to answer: this would be question-begging.

Alternatively, we could accept (RC₂) on the basis that a social welfare net is just, or good. Thus, we might say, everyone could reasonably accept a social welfare net. Once this inference is admitted (from the justice of a position to its reasonable acceptability), we have broken down the barrier between rational and public justification; but then it is again hard to see why we appeal to public justification in the first place.

4.4.4 Thin and Thick

One might reply that the distinction between minimally objective, motivational reasonableness and substantive reasonableness is gradual; there are “thinner” and “thicker” notion of reasonableness. Possibly, there is a notion of reasonableness that is *thick enough* to guarantee enough reasonable agreement, but *thin enough* to escape the criticism I outlined.

To properly assess this claim, one would need to attend to particular proposals, which I can not do here for reasons of space. However, two considerations speak in favour of a general pessimism that such a successful middle position exists. First, once we admit one constraint like (RC₂), there is enormous pressure to accept others like it. As mentioned above, we are likely to accept (RC₂) because we believe in some claim about the justice or legitimacy of a certain states of affairs (in this case, the existence of a social welfare net). But there are many similar claims which have equally strong support. We have no principled reason not to accept these other reasonable constraints (RC₃), (RC₄), ..., if we already accept (RC₂). Thus, our choice between motivational and substantive reasonableness will be much more binary than it first appears.

Second, (RC₂) strikes me as being on the level of specificity which is needed to overcome the Basic Problem. A concern for a social welfare net is central and basic for many liberals; if (RC₂) were not imposed, it is likely that reasonable disagreement would sweep away the feasibility of installing a welfare net. Thus, only substantive reasonableness which accepts something like (RC₂) is “thick enough”. But I am confident that my above criticism has shown that simultaneously this form of reasonableness is *not* “thin enough”.

Therefore, we can not solve the Basic Problem by appealing to a strongly moralised version of “could reasonably accept”.

4.5 Comparisons

Another idea, ably defended by Gerald Gaus, is that reasonable acceptance is *inherently comparative*. I will first outline some technical apparatus that Gaus uses, before I criticise his solution in two steps.

4.5.1 Optimal Eligible Sets

Gaus claims that citizens would accept the Pareto principle. According to this principle, if there is some law L that all citizens believe to be preferable to law L^* , then L^* is excluded from what Gaus calls the *optimal set*.⁵¹ Independent from this notion, Gaus introduces the idea of an *eligible set*. Laws in the eligible set are those laws which citizens unanimously hold to be better than the non-coercive baseline. If we combine these notions, we get the idea of an *optimal eligible set* of laws (Gaus 2010a, 250; 2010b, 196–7).

Gaus claims that all citizens could reasonably accept that all laws that fall outside the optimal eligible set could never be legitimate law. Call this the *Negative Claim*. Second, Gaus argues that *all* citizens could reasonably accept *all* options from the optimal eligible set as legitimate. (Let us ignore here how options are chosen from that set.) Call this the *Positive Claim*.

What guarantees that there is anything in the optimal eligible set? Gaus is making a substantial claim, the

Anti-Baseline Claim. Under normal circumstances, the eligible set is never empty – that is, there is always at least one law that individuals can agree on that is better than the no-state baseline.

In a slogan, “some law is better than no law”. Gaus appeals to a mix of Hobbesian and Kantian arguments to make this point: the state of nature is extremely bad for each individual

⁵¹ Gaus 2010a, 249; 2010b, 196. (Social choice theorists would call this the “maximal” set; cf. Gaertner 2009, 6–7.) I here follow Gaus’ more recent statement of the problem (cf. Gaus 2011, sec. 16), as opposed to earlier formulations in more epistemic language (Gaus 1996, 144–58).

(Hobbes), and coordinating on one set of laws promises great mutual gains (Kant).

A second important way in which Gaus uses the idea of the optimal eligible set is by focussing on the concept—conception distinction. We *disagree* on different *conceptions* of (e.g.) liberty, while we *agree* that liberty *in general* is a value. Thus, each citizen prefers a government that incorporates liberty in some form over a government that does not recognise liberty as a value at all.

While public reason might often be *inconclusive* – that is, it might yield a set of options among which there is no clear winner – it usually is *determinate* – that is, it clearly excludes some options and narrows the field of options we can choose from.⁵² The role of public reason is not to do away with all reasonable disagreement – that is, to resolve all inconclusiveness; what we should hope for is only that it is sufficiently determinate.

Summarise this in the

Determinacy Claim. The optimal set of laws is usually much smaller than the range of available options, because citizens agree on a wide range of *abstract* principles and values, even though they might *disagree* about *particular conceptions* of these principles and values.

If we couple Gaus' Positive Claim, Anti-Baseline Conclusion, and the Determinacy Claim, then there always is a reasonably determine optimal eligible set, and all citizens could reasonably agree that options chosen via some decision procedure from that set are legitimate. Thus, many laws deviating from the baseline are legitimate: we have avoided the Basic Problem.

4.5.2 Critique

I will grant Gaus the Anti-Baseline and Determinacy claims, which both strike me as plausible. Gaus' argument goes wrong in his *Positive Claim*.

Consider, for an analogy, the following situation: Tom stops Jerry on a highway and asks him “for his money or his life”. Assume that Tom argues:

“We would both prefer that I do not kill you. Therefore, the set of options in which I do not kill you is Pareto-superior to all those options where I kill you. Not killing you, but getting your money is part of that Pareto-superior set. Therefore, *you could agree* that you give me your money, because that is part of that Pareto-superior set.”

⁵² The distinction is drawn, with more technical detail, by Gaus (1996, 151–8), and then endorsed by Schwartzman 2004 and Williams 2000.

Tom's error is that he appeals to a comparative notion of agreement. Jerry could not agree to being robbed, *even if he could agree* to the second-order claim that, comparatively, being robbed but not killed is better than being killed.

As a second, morally less loaded example, imagine that Mary and Joseph have to decide to buy a car together. Assume that Mary argues:

"You prefer flashy sport cars, while I prefer a cheap family car. But we both think that either of these options is preferable to an SUV. Thus, you could agree with buying the cheap family car."

Mary's error in reasoning is just as grating as Tom's: she abruptly moves from a comparative claim about relative acceptability to an absolute claim. This will do little to convince Joseph.

Gaus commits the same error in his Positive Claim. We start from

(1) System of law⁵³ *X* is better than system of law *Y*.
(All citizens could reasonably accept *X* as *better than Y*.)

and from there concludes that

(2) *X* is good. (All citizens could reasonably accept *X* as legitimate.)

But this inference is false.⁵⁴ There is also a further error, as can be seen in my parenthetical reconstructions: (1) is a claim about the (*relative*) *quality* of (two) options, but we are concerned with the (binary) *permissibility* of (one) option in (2).

Gaus has a response to this objection: he emphasises that political choice is always choice from a set of options (2011, 268–70). In the process of buying a car, it would be strange to say that I have sufficient reason to choose this car without knowing alternatives and their relative prices and features. Similar, in political matters we establish that some law or institutional arrangement is acceptable by showing that it is *more* acceptable than *other* feasible laws or arrangements. The idea of some option being reasonably acceptable *simpliciter* is non-sensical. All reasonable acceptance is comparative.

Is it? Remember first that the relevant voluntaristic attitude in (JL) is accepting something *as legitimate*. Gaus's claim has some plausibility when it comes to justice. To some it will sound

⁵³ Aggregation problems surface here (cf. sec. 2.6.1), as observed in Lister 2010b. To circumvent such problems, I will speak of "system of laws" throughout this section.

⁵⁴ The inference in the Negative Claim is similarly mistaken. It is possible that individuals could reasonably accept *X* as legitimate, *even though* they also accept it as worse than *Y*. Note also that the inference from (1) to (2) is mistaken for a more general class of predicates, such as "more red" and "red", or "larger" and "large".

strange to say that institutional arrangement *X* is just when there is a wide range of other feasible institutional arrangements which strongly surpass *X* in positive qualities. Personally, I do not even find that claim odd, but let us put the matter aside.

Legitimacy is a different matter. Assume that a libertarian knows that an institutional arrangement *A* guarantees strong negative rights to everyone including property rights, and that the government upholds the rule of law in *A*, and provides security and freedom to some degree. Inversely, assume that another institutional arrangement *B* is characterised by general interference of the government with individual's lives; wide-spread disrespect for individual (property) rights; and a general breakdown in public order and peace.

I find it intuitive to claim that no comparison is needed here for the libertarian to classify the first arrangement as legitimate, and the second as illegitimate. Similar claims can be made for other political positions. In general, legitimacy is not a maximising matter. If matters are "good enough" – if there is enough concern for individual rights and interests, democratic procedures, the rule of law etc., then a state is legitimate.

We might object that under some circumstances comparisons *do* matter. It might be that we live under extra-ordinary, catastrophic circumstances, and *B* is *the best we can do*. Then, *B* would become acceptable even to the libertarian, and his acceptance is thus shown to be comparative. We can admit that sometimes, if the range of feasible options is very restricted or very unusual, comparative matters come to the fore. However, that does little to show the general point that reasonable acceptance is *usually* comparative.

Another reply could be that some non-libertarian positions are comparative. Egalitarians, for example, can judge various distributive schemes only by comparing them to the full range of feasible alternatives. The legitimacy of a distributive scheme will then depend on how far that scheme approaches the most feasible equal distribution. But what this at most shows is that such positions are *partially* comparative. An egalitarian will also be concerned with rights and liberties, equality of opportunity, offices open to all, non-discrimination and the rule of law. These can be largely judged on a non-comparative basis, and thus the overall reasonable acceptance of an egalitarian will still be largely non-comparative. Furthermore, even judging various distributive schemes does not strike me as fully comparative. For egalitarian legitimacy, it will often suffice that a distribution is "equal enough".

In summary, there is no reason to suppose that reasonable acceptance is primarily comparative. Citizens will reasonably reject some options that fall into the optimal eligible set of institutional arrangements, and reasonably accept some that fall

outside it, because the inference in Gaus' Positive Claim is mistaken. Thus, Gaus' position can not solve the Basic Problem.

4.5.3 Circumstances of Politics

There is a further argument to establish the comparative nature of reasonable acceptance, by pointing to the *circumstances of politics*. It is an important practical requirement that we coordinate on one set of laws; not everyone can follow his own conception of what the law is. While agreement on one religious creed is not necessary, agreement on one set of laws is. Thus, we might argue, once we narrowed down the set of options to those that are unanimously agreed to be comparatively best, we ought to agree to one of them; refusing agreement implies that you are unreasonable, and do not understand the specific nature of the political.

Return to Joseph and Mary. Consider a revision of Mary's argument:

“We *must* buy a car, because we need it for commuting. This is a problem we *must* solve *together*, as we are married and have a limited budget. We agree that all other options are worse than the sports car and the family car – further agreement will not be forthcoming. Refusing any agreement on the ‘car question’ would be unreasonable. Thus, you could reasonably accept that we buy the family car.”

A political argument that invokes the circumstances of politics is similar; it emphasises the fact that “we must find agreement”, and then proceeds to argue that this shows that any comparatively good option is reasonably acceptable to everyone.

This argument fails. It is true that it would be unreasonable to reject coordinating on one set of laws, or buying the family car; but that does not show that that set of laws, or the utility vehicle, is reasonably acceptable *in the right sense*. Let me explain.

We should first observe that reasons that invoke the necessity of finding a consensus do not speak in favour of any *particular* solution. They are reasons that equally speak in favour of *any* solution to the coordination problem. Mary's reasoning could be repeated by Joseph, but in favour of the sports car. Egalitarians, libertarians, perfectionists and utilitarians can all invoke the “circumstances of politics” to argue that *their* favoured set of policies could be reasonably accepted by everyone. Furthermore, if the current reasoning is correct, the libertarian could reasonably accept egalitarian policies, because egalitarian policies would be *a* solution to the coordination problem in law.

There is a great difference between arguing that “you can accept this law because it has these features” and saying “you can accept this law because we need to have some law”. In the former,

I justify *that particular law* to you; in the latter, I justify *the political relationship between us* to you.

There is a difference in the meaning of “accept” between these two modes. In the second mode, I attempt to show that *you have a reason to obey the law*, and that you ought not to undermine the exercise of power, or to interfere with it. This, again, is because everyone acting on his own set of laws would lead to chaos, social unrest, and generally decreased life chances for everyone. This is the spectre of the “state of nature” that classic social contract theorists worried about.

Thus, *obedience* or *non-interference* with the *given* law and authority is something I owe to my fellow citizens (given that the laws are not extremely unjust). In the second mode, we focus on pragmatic reasons relating to a duty to obey or not to interfere. More generally, when we say that citizens ought to accept a law or authority in this sense, we mean that citizens ought to act in certain ways.

However, the first mode of justifying *a particular law* – the mode of justification relevant in justificatory liberalism – has an entirely different focus. We justify a law (or institutional arrangement) to a person by focussing on the particular features of that law. Furthermore, our primary question is that of the permissibility of certain government actions. How individuals ought to act is secondary. The question is whether an individual *could endorse* the permissibility of the actions of some authority – that is, of some other agent; how that individual proceeds to act in the face of that fact is a different question.

Assume that a libertarian lives in a strongly egalitarian state. He could not reasonably accept, and does not accept, the high taxation he suffers from as legitimate: nothing justifies or could justify these policies to him. However, this is compatible with saying that he *could* reasonably accept, and maybe even *ought to* accept, that he has no right to unilaterally undermine this system, as long as it is not so gravely unjust that disobedience is required. In other words, individuals could reasonably accept to *confirm* with a law, and they could reasonably accept not to undermine and interfere with a given authority, *even though* they might believe the law and authority to be illegitimate. The former reasons stem from the circumstances of politics, whereas the latter have to do with the particular nature and features of the given authority.

In summary, the circumstances of politics show that citizens could reasonably accept some institutional arrangements as a unique solution to the political coordination problem, and it would indeed often be unreasonable of individuals not to accept given institutional arrangements in this sense. But this acceptance is not the voluntaristic attitude needed in justificatory liberalism; rather, this is an *ex ante*, decision-oriented, pragmatic, duty-to-obey based attitude. Thus, the argument from the

circumstances fails to show reasonable acceptance in the relevant sense to be comparative, and thus fails to support Gaus' argument.

5 The Weakest Link

I have defended the claim that justificatory liberalism falls prey to the Basic Problem. But maybe we ought to drop one of the other two claims that produces the Basic Problem – reasonable disagreement or anti-anarchism.

To clinch the case against (JL), I will use the *Weakest Link Argument*:

(WL₁) Respect for persons is a central moral commitment of liberalism. (*defended in section 5.1*)

(WL₂) Respect for persons does not require (JL). (*sec. 5.2*)

(WL₃) There are other moral principles that might lead us to accept (JL), but none of them are of central importance to liberalism. (*sec. 5.3*)

(WL₄) (RD) and (AA) are central commitments of liberalism. (*secs. 5.4 and 5.5*)

Thus, if (JL), (RD) and (AA) are collectively incompatible, as liberals we should give up (JL).

The result of the Basic Problem Argument, of course, was that (JL), (RD) and (AA) are collectively incompatible. Thus, the current argument achieves the aim of this thesis: we should give up justificatory liberalism. I will now discuss this argument in detail.

5.1 Liberalism and Respect

Let me start by defending (WL₁), the claim that liberalism is based on a notion of (equal) respect. I will first outline what we should understand by respect, before arguing that respect for persons captures the core commitment of any liberal view.

5.1.1 Respect for Persons

The notion of respect is contested, and I do not claim in the following to capture all possible interpretations of that idea. I focus on a specifically *liberal* and *political* notion of respect. First, the relevant notion of respect is guided by moral concerns: we can have respect for the power of a tsunami, or the ruthlessness of a tyrant. Both these forms of respect are rationally mandated; but they rest on prudential features, which are not of particular interest here.

Second, I focus on what Darwall has coined *recognition respect* (Darwall 1977). When we have recognition respect towards some person, our respect is based on some feature or property that person possesses, and our *behaviour* is guided or con-

strained in certain ways by this recognition. This is unlike appraisal respect, which merely consists in some positive evaluative *attitude* towards qualities that persons possess.

We can entirely ignore respect as an *attitude*, even though this is an important part of the meaning of respect. We are interested in coercion. While the attitude with which persons coerce is sometimes morally relevant, we will be largely able to explain the wrongness or rightness of coercive acts without reference to the respect-based motives of the coercers.

Every conception of respect, to be reasonably determinate, will need to clarify two aspects: first, on which particular features of human beings respect is based (the *respect-demanding features*), and second, how respect for these particular features translates into moral permissions and demands (the *intermediate premises*). Schematically, an argument from respect for persons proceeds as follows:

- (1) Citizens possess F-features which are respect-demanding.
- (2) Respect for persons with F-features requires/permits us to φ in S-situations.⁵⁵
- (3) Thus, we are required/permitted to φ in S-situations.

Let me briefly make some remarks on (1). Kant urged us to always show respect for “humanity” in ourselves and others. This has led to a long-standing debate what exactly Kant meant by that phrase: the capacity of human beings to have a good will? the noumenal part of human beings? actual human beings, with their actual desires and wishes?

It is no accident that Kant failed to give a clear answer, as different answers lead to different ethical outlooks. Assume that we see persons primarily as a bundle of desires and preferences. Then, respect for persons is most likely to entail a concern for their well-being, and the satisfaction of their desires.⁵⁶ But we could also highlight two other features of human beings: their capacity for *autonomous choice*, and their existence as *rational* animals, i.e., as those animals that are capable of giving and understanding reasons, and being motivated by them. In such a position, we are likely to draw different inferences from respect for persons, as we will see below.

Liberals demand *equal* respect for persons. This means in particular that persons are due the same amount of respect even

⁵⁵ Cf. Wood 1999, 174. We could also call these “specificatory premises”; cf. Nickel 1982, 79.

⁵⁶ Dworkin draws a distinction between “equal respect” and “equal concern” (1981, 273), where the latter is focussed on well-being. In my usage, however, “respect for person” is the genus to which Dworkin’s equal respect and equal concern belong as species.

though their endowment of the respect-demanding features might differ. We should also not confuse equal respect with equal *treatment*; while the former might often imply the latter, this is not necessarily the case.

With regard to (2), I can not here assuage worries that respect for persons is a gravely indeterminate (if not empty) concept from which we will be unable to derive particular duties, other than very trivial ones.⁵⁷ The very vagueness of the concept, however, has two advantages in the current context: first, remember that we are interested in respect for persons as the core commitment of liberalism. Vagueness is a benefit here, as we do not want to ascribe to liberalism an overly narrow moral basis, but one that is compatible with the full range of liberal positions.

Second, it might turn out that the concept is too indeterminate to derive (JL) from it. In this case, we should say that (JL) is *not* a core commitment of liberalism, even though *compatible* with it. Thus, the burden of proof is shifted onto the justificatory liberal: in the absence of a determinate argument, (WL₂) should be our baseline assumption.

5.1.2 Respect as a Foundation

It is impossible to give a deductive argument that shows that liberalism, a rich and vaguely defined historical tradition, is based on respect as persons, itself a broad and vague ideal. I also do not claim that respect for persons *uniquely* identifies liberalism; it might be the foundation of other positions as well, such as conservatism. Instead, I will try to make it plausible through four arguments that a *very broad range* of liberalisms are *compatible* with conceiving respect for persons as their central moral commitment.

First, there are the historical origins of liberalism. Liberalism first arose in reaction to views that saw God's will, or super-human aims as the centrally important concepts in political philosophy. Classic liberals like Hobbes and Kant stressed against these tendencies that political philosophy should start with the individual, and the rights and liberties due to it. This claim appears trite and trivial to us nowadays only because it has been so successful in convincing nearly all of us.

Second, the notion of respect ties together well many classic liberal concerns. At its most basic, equal respect for persons expresses the idea that individuals are *sources of valid moral claims on us*, or more drastically, that they are *sources of value*. The lives and demands of persons matter to us.

This basic idea branches into a number of further liberal principles: for example, that we should take serious the choices that

⁵⁷ For various critiques with different emphases, see Nickel 1982; Frankena 1986; Rawls 1999b, 513; Raz 2001, ch. 4; Neumann 2004; Parfit 2011, 233–5.

individuals make, and that we should not ride roughshod over people's concerns and beliefs. Respect can also lead to the claim that it is always impermissible to treat individuals in certain ways, e.g., torturing them for fun. Thus, (human) rights express respect for persons by shielding persons from certain forms of interference with their lives which can never be considered disrespectful.

On the same basis, we can explain the presumption against coercion, with which I started this thesis. Coercion always involves thwarting another person's will. As we have great respect for individuals exercising their will on their own, this directly gives us strong *prima facie* reasons not to coerce them. And lastly, a predilection for democracy can be explained on the grounds that giving some more voice in the political process than others expresses a basic disrespect for those with a lesser voice (cf. Christiano 2008).

Third, many liberals have explicitly based their liberalism on equal respect, in particular justificatory liberals. Charles Larmore has argued for a form of political liberalism that uses the idea of equal respect as its ultimate foundation (Larmore 1990; 1999). Christopher Eberle lists the liberalisms of Galston, Gutmann and Thompson, Solum, Bird, Gaus, Macedo, Rawls and others to be explicitly or implicitly based on the idea of respect for persons (Eberle 2002, 53–4, 351). Eberle also highlights that all these philosophers accept some inference from equal respect to (JL).

As a side note, this implies that even if (WL₁) were not true, my argument would at least be a good *ad hominem* attack: as the overwhelming majority of justificatory liberals accept equal respect as a ground, showing that this ground does not imply (JL) would be a major problem for their views.

Fourth, what shall we say about those liberalisms that do not use the idea of equal respect? Two other promising values which have been invoked as the foundational values of liberalism are equality and liberty. While these approaches differ in their rhetorical approach, they are entirely compatible with the idea that liberalism is grounded on equal respect for persons.

Ronald Dworkin, for example, grounds his egalitarian liberalism on an equal right to concern and respect (Dworkin 1981, 180–3). Dworkin agrees that this right is more basic than particular conceptions of equality, and he claims that this deeper concept enables us to decide between different articulations of equality.

Jeremy Waldron, while stressing the importance of liberty as central to liberalism, also writes that liberals “are committed to a conception of freedom and of respect for the capacities and the agency of individual men and women, and [...] these commitments generate a requirement that all aspects of the social should either be made acceptable or be capable of being made acceptable to every last individual” (1987, 128).

In general, “equality” and “liberty” strike me as good, if competing, characterisations of the *content* of liberalism. But in both cases we can ask “why equality?” and “why liberty?”. And the most plausible answer – at least, a *compatible* answer – is equal respect for persons. Thus, while I do not want to deny that there are competing rhetorical representations of the liberal core commitments, none of them contradict the claim that respect for persons forms the core of liberalism.

5.2 Respect and Justification

What about (WL₂), the claim that respect does not require public justification? Christopher Eberle has analysed in-depth and rejected a number of particular approaches that attempt to derive (JL) from equal respect for persons – arguments by Solum, Larmore, Audi and Gaus.⁵⁸ I am convinced that Eberle’s critiques are successful, and I do not want to rehearse his discussion here.

Eberle’s project is largely negative, in that he knocks down existing arguments from equal respect to (JL). In contrast to that approach, I want to make a more general argument why derivations of (JL) from equal respect for persons fail.

5.2.1 External and Internal Conflicts

A first argument is as follows: respect for persons is not the only relevant moral (liberal) principle or value. At the very least, there is the consequentialist principle of maximising the aggregate good. In some cases, these other principles will override the value of respect for persons; thus, there will be some legitimate (maybe even required) policies where respect for persons has to be set aside – and with respect for persons, justification to all citizens, if we assume that the former implies the latter. Thus, we ought not to pursue public justification for all laws.

Liberals could reply that respect for persons is indeed the *only* moral principle that should guide us in making law. This would be a very strong reply. Instead, liberals can say that respect for persons is not a principle or value that is *weighed against* other liberal principles and values, but works as a *meta-constraint* on the pursuit of these other values. Respect modifies *how* other values can be achieved, but does not make their pursuit impossible. This is still not an innocuous assumption, and would need further defence. But let us assume that this position can be defended.

⁵⁸ Eberle 2002, chs. 4–5. Other defenses of (WL₂) can be found in Galston 1991, 108–9; Wall 1996, ch. 4; Wolterstorff 1997, 109–111; McCabe 2000, 324–8; Arneson 2003, 211–4. See also the papers by Reidy (2007) and May (2009) in a Rawlsian context.

A second argument proceeds from conflicts *internal* to the idea of respect for persons.⁵⁹ Assume that we have a good argument that the existence of a certain form of human rights is necessitated by respect for persons. At the same time, assume that these human rights could not reasonably be accepted by all citizens. Insofar as we believe that the inference to the existence of human rights is stronger, we would have to deny (JL); we have to follow what respect for persons requires, not what others can accept, even if following the latter is implied by respect for persons as well.

The same argument can be repeated with a different focus – some commentators have derived the claim that democracy, or strong egalitarian principles are required by equal respect for persons (see sec. 5.1.2). Similarly, these commitments might be reasonably rejected by some. Internal conflicts in the notion of respect are not surprising at all. If you are deciding to take an addictive drug, I have respect-based reasons which go both ways: to respect your free choice (and let you do it), and to respect your future well-being and autonomy (and prevent you from doing it).

We might argue, similar to the reply chosen above, that whatever respect for persons demands of us (e.g., human rights), it is a side-constraint or meta-demand that all implications of respect for persons be reasonably acceptable to everyone. Thus, insofar as human rights could not be reasonably accepted by everyone, they are not required by respect for persons. But this argument is unconvincing in its new context. What is special about the inference we draw from respect to reasonable acceptance that sets it apart and makes it such a meta-demand?

An alternative reply is to say that the different implications of respect for persons do *not* conflict in the mentioned way. We can appeal to a morally substantive notion of reasonableness and argue that all citizens could *reasonably* accept the particular implications of respect for persons. But this response will only carry us so far. The more specific the implications of respect for persons become, the more implausible a moralised reading of reasonable becomes (see sec. 4.4).

In other words, justificatory liberals either have to deny that there are internal tensions to the idea of respect for persons, or they always have to resolve internal tensions in favour of the “procedural” side of equal respect (public justification), while subordinating the “substantive” side (e.g., human rights and distributive concerns). Both solutions are difficult choices. The first solution relies on a mysterious pre-established harmony between the procedural and the substantive side of equal respect. The second solution is counter-intuitive, because we do

⁵⁹ For similar remarks in the context of the duty of civility, see Lott 2006. Cf. Lister 2010b, 168–9.

not generally choose procedural over substantive implications of respect for persons.

These general and somewhat global preliminary points can do well to lead us into the main critique of the inference from equal respect to (JL). As we will see, the argument for justificatory liberalism only succeeds by implausibly pushing many values aside. But this is something liberalism should generally not be committed to.

5.2.2 The Reverse Strategy

My main argument relies on the *reverse strategy*. Our argument from respect for persons in the current context must proceed as follows (see above):

- (1) Citizens possess F-features which are respect-demanding.
- (2) Respect for persons with F-features requires us not to coerce them in ways that they could not reasonably accept.
- (3) We are required not to coerce citizens in ways that they could not reasonably accept. (This is (JL).)

The reverse strategy looks at what F-features are most likely to yield the conclusion. I will discuss two such features in the following: (a) individuals' capacity for reason, and (b) individuals' capacity for autonomy. I will argue that neither of these grounds makes premise (2) likely.

There are other F-features which could be considered. However, the two I mentioned are the ones which are most closely connected in their central notions – reasoning and willing – with (JL). Respect for individual's capacity for well-being, for example, is only indirectly connected to (JL). It is often the case that being subject to a policy I could not reasonably accept negatively influences my well-being. But this link is highly contingent and empirical, and there are many ways in which a government can respect individuals' well-being, even if this is through policies some citizens could not reasonably accept. This makes it unlikely that (JL) can be derived from respect for individuals' well-being. Generally, I assume that if the arguments from respect for reason and from respect for autonomy fail, then other arguments from respect fail as well.

5.2.3 Respect for Reasoners

Let me first consider individuals as reasoners. From previous discussion, we know that “respect for reasoners” is ambiguous between respecting individuals' capacity for *theoretical* reason, and their capacity for *practical* reason. Respect for reasoners is likely to extend to both.

What does it mean to respect someone's *capacity* for x ? (x could be, for example, creativity, self-guidance, athleticism, mathematics etc.) It means, for example, that we should not interfere with others' x -ing; that we should encourage others in x -ing, and help them to improve how they do it; and to provide resources and freedoms for others to exercise their x -ing.

If we fill in x as "reason", a number of duties immediately suggest themselves: we ought to provide education to all citizens, and we ought not to contribute to the "stupefying" of citizens. Generally, governments should guarantee circumstances under which the full exercise of one's theoretical and practical reason becomes possible. This will also impose a duty on governments to publicly give (good) reasons for coercion, and have a decision-making process that is guided by reason and evidence. (This is the first sense of "public justification" discussed in section 2.2.)

However, these policies are all compatible with treating people in ways that they could not reasonably accept. The pursuit of social justice, for example, might not be reasonably acceptable to some citizens; but it does not necessarily interfere with their capacity to reason fully and freely. In fact, some policies that the majority of people could not reasonably accept might increase or better preserve their ability to reason, such as forced education. We might be morally *required* under the current argument to treat citizens in ways that they could not reasonably accept.

Indeed, some commentators have claimed that respect for reasoners requires us to treat others in the most *objectively* reasonable way, not in ways that the coerced individuals *perceive* as the most reasonable.⁶⁰ According to these critics, I respect that others are capable of reason *only* by giving the *best* available practical and theoretical argument for my position. I should appeal to the truth, and nothing but the truth; if I were to give heed to what I see as flawed reasoning in others, I would treat them *as unreasonable*, rather than reasonable.

Consider an analogy. If I stand under a duty to respect the capacity of someone for mathematical reasoning, this does not mean that I have to accept her faulty proofs and calculations. It implies that, if she makes a mistake, I should do my best to try and correct her, show her the correct proof or calculation, and to help her to avoid making such mistakes in the future. My respect for her (mathematical) reasoning *capacities* does not translate into respect for her (mathematical) *beliefs* that are the *outcomes* of the *exercise* of that capacity.

Here we encounter an important distinction. There is a difference between respecting a *capacity* for x and respecting the *exercise* of a capacity for x . If I only ought to respect Peter's *capac-*

⁶⁰ Galston 1991, 108–9; Wall 2010b, 253–4. Wall also claims Bohman and Richardson (2009) as supporting this position.

ity for autonomy, I ought not to let him take an addictive drug, which will inhibit his autonomous capacities in the future. If I only ought to respect Peter's actual *exercise* of autonomy, I ought to let him take it, given that he chooses to do so freely.⁶¹

Insofar as we only ought to respect the *capacity* for reason in individuals, we will be unable to derive (JL). On the other hand, if we ought to respect individuals' actual exercise of reason, it will be more plausible that respect for reasoners translates into respect for the beliefs these persons have. If you reason your way to a libertarian position, then I ought to take libertarianism into account when formulating policies; not because it is true, but because *you* believe it, and my prime duty is to respect *you*, which translates into respect for *what* you believe.

Charitably interpreted, many authors who advocate respect for the reasoning capacities of others probably meant to say that we ought to respect some mix of the capacities themselves *and* their exercise. However, it is helpful here to focus on the two analytically pure versions, and show that both of them fail. If respect for the capacity itself does not imply (JL), and respect for its exercise does not imply (JL), then no mixed position will entail (JL) either.

Thus, we now have to discuss respect for the *exercise* of reason. On closer reflection, however, I think that we do not really respect the actual exercise of reason *per se*, as can be seen in the example of mathematical reasoning given above. Considered as such, we value the exercise of reasoning only insofar as it leads to true results. If we value it, this is primarily due to other values and capacities being valuable. The primary value here is *autonomy*. We ought to respect the reasoning of others because we ought to respect that others reason *autonomously*. Thus, respect for the exercise of reasoning centrally depends on respect for autonomy, to which I now turn.

5.2.4 Respect for Autonomous Choosers, I

There is a vast literature on autonomy, which I can not adequately tackle here.⁶² Intuitively, personal autonomy is our ability to make free, self-guided and unconstrained choices. Again, we can distinguish a capacity for autonomy from the exercise of that capacity (cf. Feinberg 1989). An argument similar to the one in the last subsection could be made to the effect that respect for the *mere* capacity of autonomy is unlikely to yield (JL): treating you in ways you could not reasonably accept is often compatible with you retaining your capacity to choose autonomously in the future.

⁶¹ There is a further difference between respect as *passive non-interference* with a capacity or its exercise and respect as *active assistance*, which might overlap with the current distinction.

⁶² A good overview on many of the distinctions I discuss here can be gained through Christman 1988; 2011; Christman and Anderson 2005; Buss 2008.

Given that we ought to respect the actual exercise of people's capacity for free and self-guided choice, are we committed to (JL)? Not on all readings of autonomy. At least on a major family of views, individuals choose autonomously only if their choices are appropriately sensitive to what (objective) reasons there are. This family of positions has been called "substantive", and more tellingly, "externalist".

The strongest classic advocate of this reading is Kant, for whom we are autonomous only insofar as we choose the moral law.⁶³ On this strong reading, it is clear that respect for autonomy does not require us to treat people in ways they could reasonably accept, where we understand that phrase in an internalist sense. Your beliefs in moral propositions that are not in accord with the moral law are not autonomous, and as such not deserving of respect.

Modern defenders of substantive interpretations of autonomy advocate more moderately externalist approaches. However, their approaches still lead to similar results: the externalist interpretation of autonomy means that (JL), where we read "could reasonably accept" in an internalist sense, is not implied by a principle of respect for autonomous choosers, where autonomy is subject to externalist conditions.⁶⁴

This gives us a first argument for (WL₂). The current position shows that there can be a liberalism that accepts the moral core commitment of liberalism – respect for persons understood as respect for autonomy – but that does *not* accept (JL). Thus, (JL) is only a contingent commitment of liberalism, and not a moral core commitment.

Justificatory liberals might reply that externalist readings of autonomy are *wrong*. But that would not be enough. They would have to show that such interpretations of autonomy are not even *prima facie plausible* as *liberal* commitments. But such an argument is extremely unlikely to be available. Externalist conceptions of autonomy, whatever else you think about them, are not *obviously* wrong or implausible, and they are serious contenders in the debate about autonomy.

Second, it is unclear why liberalism would be committed to reject such conceptions of autonomy. The debate about autonomy reaches deep into meta-ethical and meta-physical terrain. Liberalism, if defined broadly, ought not to be committed to very specific claims on this level; and if we speak of liberalism as a historic and broad tradition, it is wrong to think that it is committed in any such way. Thus, even if justificatory liberals are convinced that substantive readings of autonomy are *wrong* (and this must be their conviction), there still is a wide family of

⁶³ This is a very simplistic picture; Kant's position is more complicated than this (cf. Waldron 2005).

⁶⁴ It is no accident that prominent proponents of externalist readings of autonomy (Sher, Wall, Hurka) are at the same time political perfectionists.

plausible respect-for-autonomy-based liberalisms that do not accept (JL). But then (WL₂) is wrong.

5.2.5 Respect for Autonomous Choosers, II

I will now give two further arguments for (WL₂), by turning to “procedural” (“coherentist”, “internalist”) interpretations of autonomy. According to these views, our choices count as autonomous insofar as they were chosen in a certain way, independent of *what* we have chosen. For example, my choice counts as autonomous if it fits into a “coherent narrative” of my life, or if I can reflectively endorse my choices, possibly through “second-order” desires and beliefs.

Thus, if we are to respect the autonomous choices of others, then we ought not to act in ways that make it impossible for individuals to reflectively endorse their own choices and desires.

1. Note that there is a gap between coercion and autonomy-infringing interferences. Coercion *usually* disrupts people’s ability for autonomous choosing, but the connection is contingent and empirical. Coercion does not make it generally impossible to reflectively endorse or narratively acknowledge what happens to oneself, as can be seen in the extreme example of the slave who has come to reflectively endorse his own enslavement. Inversely, coercion is not even *necessary* to infringe upon an individual’s autonomous choosing. I might undermine your sense of narrative self-identity without coercing you in any way.

These empirical connections are also subject to individual and cultural variance. In some cultures, not being able to choose one’s spouse is not seen as a great infringement on one’s ability to lead a self-guided life. Furthermore, the infringements on autonomous choice come in degrees. Forcing a religion on you deeply disrupts your autonomous choosing; forcing you to wait in queues in front of busses does so to a much lesser degree, if at all.

If we follow the argument, we can see that the areas of human life which will require the most protection are those where state interference has an especially strong negative impact on autonomy, understood as reflectively endorsed choosing. In particular, there will be strong protection for the “inner citadel” (Christman 1988) of individuals’ lives – the area where individuals choose their own identities, and where they develop the set of preferences, wants, and desires which they think defines them. Thus, this argument would naturally lead us to accept a strongly protected area of privacy, and a set of rights and freedoms that protects individuals’ choosing of their core identities. For example, we ought not to force individuals to take up occupations, religions or homes against their will.

2. But this does not give the result that only laws that everyone could reasonably accept are legitimate, and it is unlikely that such a result will be forthcoming. In connection with the in-

junction to respect others, we should remember Kant's famous formula that we should treat others "not merely as a means". The emphasis here is on "not *merely*". A government can use citizens as a means – it can treat citizens in a way that they could not reasonably accept – if it *also* respects them as an end in other ways.

One important way to treat individuals as an end is to confer some benefit on them. Many liberals have thought that the fact that citizens are subject to coercion imposes a duty on the state to provide everyone with some level of distributive goods.⁶⁵ This is a weighing question, of course; the fact that the state provides benefits to citizens will not *always* make it true that it does not treat them merely as a means. But sometimes it will. A political minority might be subject to political decisions they could not reasonably accept; but if the government protects the mentioned "inner citadel", and provides some benefit to these individuals, then the government will often not treat them merely as a means. This will be the case even if the minority thinks that the provided benefits do not make the overall treatment they receive reasonably acceptable.

Justificatory liberalism based solely on respect for autonomous choosers is committed to the claim that providing the political minority with benefits would never be a way of treating them as ends. On the current position, *only* if I react to your autonomous choices do I respect you as an end.

This is implausible. There are several ways to treat you as an end. It is correct that respecting your autonomous choosing is *one* such way, and the current argument can be used to derive protections for the "inner citadel". But the argument goes astray in supposing that this is the *only* way.

We can couple this back to the standard three-step argument from respect (sec. 5.2.2). Citizens bear *several features* which are respect-demanding; the exercise of a procedurally understood capacity for autonomy, even if important, is not the only one. In the current context, the fact that individuals have *interests* and are capable of well-being is one such other feature which is respect-demanding. This is an obvious feature of human beings, and one which any plausible political anthropology has to take into account; but justificatory liberalism fails to do so. Thus, there are other ways to respect others as ends; and the current argument suffers from an impoverished view of human beings. This again suggests (WL2).

3. However, let us admit that *only* the exercise of procedural autonomy demands respect. Even then, it does *not* follow that it is never permissible to treat others in ways that they could not reasonably accept. If in some matters I treat you in ways that

⁶⁵ This is an influential claim in the context of debates about cosmopolitanism. See for example Blake 2001; Nagel 2005.

you could *not* reasonably accept, and in other matters I treat you in ways that you could reasonably accept, it is still possible that I treat you as an end – exactly because I do not treat you *merely* as a means.

For example, assume that there are a few libertarians in a state dominated by egalitarians. The libertarians will foreseeably be the group that most disagrees with many policies of the (egalitarian) government. What does respect for autonomous choices require of the egalitarians? It requires guaranteeing the set of “privacy” and “personal” rights and liberties that I mentioned above. Furthermore, the state might install relatively strong property rights, will allow citizens to opt out of certain distributive schemes (e.g., compulsory education, the public health system etc.). This goes a far way to respect the libertarians’ individual choices.

However, these accommodations are *not* intended to make the libertarians *agree* with the coercive order, or make it the case that they *could* reasonably agree. The egalitarian government will pursue a wide variety of egalitarian policies, which the libertarians could not reasonably accept. For example, the state might pursue affirmative action, or install an extensive system of social welfare.

Respect for autonomous choosers, and respect for persons more generally, requires us to be sensitive to the autonomous choices made by citizens. But it does not require us to be sensitive to *all* of these choices. It requires us to respect central and important autonomous choices, – especially those belonging to the “inner citadel” – but it is compatible with treating the individual in other ways that it could not reasonably accept.

In summary, even on the very strongest interpretation of respect for persons, respect for the exercise of procedural autonomy, and even if we implausibly make this form of respect the sole relevant form of respect, (JL) does not follow.

I take these two arguments to establish (WL₂), together with the independent arguments in subsections 5.2.3 and 5.2.4.

5.3 Other Arguments

We might defend that (JL) is a central moral commitment of liberalism in other ways, which I now discuss.

First, we might say that (JL) simply *is* a central moral commitment of liberalism. This is not uncommon; “neutrality” between conceptions of the good, for example, is sometimes cited as a defining feature of liberalism. But in the current context, such a claim is simply question-begging; we want to know *why* (JL) is such a commitment.

Certainly, (JL) has not historically been part of liberalism, and we can not find it in classical formulations such as Locke’s, Mill’s, Constant’s or Humboldt’s (Grotefeld 2006, 250–268).

Kant's liberalism might seem the exception to the rule; but close analysis would reveal, I believe, that Kant's language in *On Theory and Practice* does not actually suggest (JL), but rather a duty to *rationally* (not publicly) justify policies.

Let me shortly discuss (and dismiss) three other arguments in favour of (JL) being a central commitment of liberalism.

5.3.1 Meta-Ethical Arguments

One tempting argument might be this:

- (1) Coercion requires giving moral reasons.
- (2) Moral reasons require, constitutively, public justification.
- (3) Thus, coercion requires public justification. (This is (JL).)

Jürgen Habermas can be interpreted to advance such an argument. Habermas claims that it is constitutive of making a moral claim that one thinks that the norms advanced could be reasonably accepted by everyone (Habermas 2004). Thus, the nature of moral argument commits us to public justification. While very different in its details, Gerald Gaus also increasingly sympathises with such a meta-ethical line of argument.

We should reject such an argument for three reasons. First, for (2) to be a plausible meta-ethical claim, we would need to read reasonable acceptance through a heavily externalist standard. Habermas claims, for example, that the validity of moral norms depends on their acceptability in an *ideal* speech situation. But what individuals accept in such a situation will deviate strongly from what they actually accept, or could reasonably accept on an internalist standard. Generally, meta-ethical theories that try to explain ethical facts through “ideal observers” or “perfectly rational deliberators” utilise very heavy idealisations of actual people; but then this is unlikely to yield (JL), which focusses on what people could reasonably accept under an internalism constraint.

Second, the current argument would fail to establish justificatory liberalism as a convincing *normative* position. Assume that a meta-ethically grounded justificatory liberal attempts to publicly justify his coercive demands towards others. He does not really *care for others*; the opinions of others *are mere evidence for moral truth to him*. But public justification is an essentially directed phenomenon: it is justification *to* people. This expresses that something about people fundamentally matters; but the current argument is unable to explain this.

Third, even if the last two problems can be avoided, the current argument fails to establish (JL) as a moral *core* commitment of liberalism. The meta-ethical views which would underlie (2) will be very controversial. But it would be strange if a commitment

to (say) Habermasian meta-ethics would be a *constitutive* commitment of liberalism, especially as this is a claim in *meta*-ethics. Liberalism should be neutral in its foundations between realism and anti-realism, and other meta-ethical views.

5.3.2 Democratic Theory

Another form of argument might proceed from democratic theory.⁶⁶ In a democracy we all equally ought to be *authors of the law*, or on a weaker reading, we should all equally be capable of seeing ourselves *as if* we were authors of the law. We could then claim that we could not see ourselves as authors of a law which we could not reasonably accept. This would imply (JL).

The fiction of all citizens being authors of the law (or of being able to regard themselves *as if* they were such authors) is a tempting liberal ideal at least since Rousseau. One might answer directly that this Rousseauian ideal was never more than a utopia – even in its “as if” formula – and should thus not be considered as a central liberal commitment. Or maybe, as an ideal, it is not a claim about real-world conditions of legitimacy, but about the ideally just form of government.

A more constructive reply would be to say that joint authorship of the law does not, in fact, require (JL). Laws can be jointly authored, and still not reasonably acceptable to everyone. What matters for the former is that everyone’s voice has been heard in the process of making law, and that everyone had equal power in that process. That we have to disregard the opinions and preferences of some in the end does not mean that the laws are not made “in their name”, and that they are not co-authors.

More generally, democratic equality requires that everyone has an equal voice in the political process; but it does not require us to make laws equally acceptable to everyone.

5.3.3 Stability

A third argument proceeds from the values of stability, public peace and cooperation. We should accept (JL), a justificatory liberal might claim, because this is the only way we can achieve these values which are of prime importance.

This argument can be read as a consequentialist argument that policies constrained by (JL) are more likely to (causally) bring about peace and the absence of social unrest, which are important pre-conditions for individuals achieving well-being, and other values. But it is unclear whether this empirical, causal connection is empirically true – that is, whether such a strong principle like (JL) is *required* to bring about social peace. We certainly know of peaceful societies that exist in the face of deep disagreements, and where governments pursue policies that do not stay neutral in the face of reasonable disagreements.

⁶⁶ For such an argument, see Bird 2011. Cf. Lister 2008, 276–7.

Second, the values mentioned do not strike me as important enough to regard (JL) as a core commitment of liberalism, even if we grant the causal claim. Assume that we base (JL) on the value of stability. Then, the Basic Problem arises again, and we have to choose between realising the value of stability, or the value of justice. But this strikes me as a decision which is very easy to make: we should decide in favour of justice, and should secure stability only insofar as an absence of stability interferes with pursuing just policies, or diminishes justice.

Similar arguments can be made with regard to other consequentially relevant values that (JL) might help bringing about. We can question them either from the causal side (it is unlikely that (JL) is the only way to bring these values about) or the normative side (it is unlikely that these values are strong enough to lie at the foundation of a requirement as strong as (JL)).

These remarks make it likely that (WL₃) is true – that is, that there is no other argument from liberal premises that establishes (JL) as a central liberal commitment.⁶⁷

5.4 Liberal Commitments: Reasonable Disagreement

Even if we are correct in supposing that (JL) is detachable from the core of liberalism – i.e., even if (WL₂) and (WL₃) turn out to be true – it is still possible that (RD) and (AA) express ideas which are connected to liberalism on an even more tenuous basis. I will now argue that this is not the case.

Let me start with (RD). Reasonable disagreement, as referred to in (RD), is defined as any situation in which one citizen could reasonably accept some law or institutional arrangement, and another citizen could not. What should be clear is that the relevant sense of “could reasonably accept” depends entirely on the sense in which it is used in (JL). It would be pointless to claim that there is no reasonable disagreement *in an externalist sense*, because this would not help to solve the Basic Problem. Instead, the relevant sense of “could reasonably accept” must exactly be the one I specified at length in chapter 2.

Once we have this in mind, (RD) should strike us as a simple fact. (RD) is neither a normative claim, nor a strongly epistemological or metaphysical claim. That we ought to expect reasonable disagreement in the internalist sense is a common-sense observation, supported by opinion polls, sociological observation, and explanations such as Rawls’s “burdens of judgment” (see sec. 3.1). Thus, I think that liberalism should accept (RD) simply insofar as (RD) is a fact, and liberalism ought to acknowledge the facts.

⁶⁷ Arguments I could not assess here are from the value of civic unity (Lister 2008, 277–9), and a recent republican argument (McCammon 2011).

We can also cite three further reasons why liberals should be especially willing to accept (RD). First, the classical liberalism of Locke, Kant and Berlin were all in different ways responses to the post-Reformation insight that we no longer live in a society with shared horizons of meaning, religion, history, or ethnicity. Liberals have classically valued, or at least taken especially serious, the fact of pluralism in human life. Thus, they should be particularly cautious when it comes to denying reasonable disagreement, or underestimating the depth and width of such disagreements.

Second, liberalism is a pragmatic view that attempts to find solutions for largely non-ideal circumstances. While there might be ideal convergence on truth, we are interested in actual political problems. And for actual politics, ignoring the fact of strong disagreement would be naive and misguided.

A third reason again stems from respect for persons. It is often disrespectful to count the dissent of others as unreasonable, especially if that dissent is not obviously unreasonable, or standards of reasonableness are themselves contested. Liberals should be cautious in discounting others as unreasonable. There are again pragmatic experiences which should make liberals wary: historically, literacy requirements have been used in the United States to disenfranchise blacks.

These three claims – the historical experience, pragmatic focus, and grounding in respect – show that liberals have especially strong reason to take reasonable disagreement seriously, beyond simply acknowledging it as an empirical fact. A liberalism that denied (RD) would be ignorant of empirical facts, oblivious to the modern historical situation, unrealistically utopian, and disrespectful to the internal perspective of agents. This makes (RD) a strong commitment that all liberals should accept.

5.5 Liberal Commitments: Anti-Anarchism

Why is liberalism centrally committed to anti-anarchism? We can imagine a position we can label “liberal anarchism”. A liberal anarchist accepts that legitimate coercion needs to be publicly justifiable. However, she sees legitimate coercion as attainable only under very rare circumstances, if at all. This is not an “a priori” anarchism, which regards state intervention as generally illegitimate or undesirable; it is an “a posteriori” anarchism which claims that legitimacy is not feasible under actual conditions (Simmons 2009).

Gerald Gaus' work can be described as partially accepting this consequence. While not an anarchist, the publicly justified state Gaus has in mind has strongly libertarian tendencies.⁶⁸

Is there anything in liberal thought that prevents liberal anarchism? The simple answer is that liberalism is committed to ideals of *social justice*. Liberalism has always been an ideal on how we should live together in a society, and how we should share the burdens of cooperation in a state. Most liberals believe, for example, that the state should engage in extensive redistribution of social goods.

We should remember that (AA) is just a placeholder for whatever substantive, non-trivial policies liberals want the state to pursue, and where they intuitively think that such pursuit is legitimate (cf. sec. 3.2). Thus, there are as many arguments that (AA) is a central commitment of liberalism as there are forms of liberalism. Such divergent liberalisms as the ones of Dworkin, Sen, Berlin, Locke, or Kant all are centrally committed to (AA), as they will all entail that some non-publicly justifiable policies are legitimate, and even required.

Thus, it is unclear whether much more can be said in a systematic fashion that (AA) is a central commitment of liberalism. We might point to the substantive side of respect for persons (sec. 5.2.1). Insofar as liberals are committed to an ideal of respect for persons, they think that various treatments and forms of government intervention are morally required, and legitimate. But this only gives a very high-level explanation of (AA).

The best we can do is to appeal to the particular moral and political intuitions of actual liberals. The claim that non-trivial policies such as the pursuit of social, distributive and historical justice and many other goods are illegitimate strikes me as an obvious *reductio ad absurdum* of any position that has such an implication.⁶⁹ Of course, liberals should be open to the suggestion that *particular* policies turn out to be illegitimate, and the absence of reasonable acceptance can play a role here. However, it is implausible that *all* liberal policies turn out to be illegitimate.

This is not a deductive argument, but an appeal to a strong set of intuitions. However, as these intuitions are nearly universally and very strongly held across the liberal spectrum, I think that we can plausibly take (AA) to be a moral core commitment of liberalism.

⁶⁸ Gaus' most explicit defense of this claim can be found in his 2010a and 2010b. Cf. Lister 2010b.

⁶⁹ E.g., Caney 1998; Raz 1998; Reidy 2007.

6 Concluding Remarks

If my argument is correct, we should abandon (JL). Instead of a summary, let me highlight a number of questions for further research.

A first set of questions concerns the status of public justification. One might try to evade the Basic Problem by accepting

Ideal Theory (IT): (JL) is true, but (JL) only applies under ideal circumstances.

In other words, governments under non-ideal circumstances do not have to publicly justify their policies to count as legitimate.⁷⁰ There is a major problem with (IT). If we assume “ideal” circumstances to be “circumstances under which the Basic Problem does not arise”, then (IT) begs the question against the Basic Problem. If reasonable disagreement *still* arises under ideal circumstances, the Basic Problem has not been solved.

A more plausible claim is

Pursuit of an Ideal (PI). It is a *valuable good* that policies are publicly justified, and we should *pursue* public justification *as far as possible*.

(PI) is a claim about what makes a government *good*, not about what makes it *legitimate*. A society in which wide-spread justification to citizens is achieved is a tempting liberal ideal, and the Basic Problem does not spell doom for the *desirability* of that ideal. The set of arguments which would be needed to refute (PI) would be quite different. This promises hope for justificatory liberals, by moving from legitimacy to justice.

A second set of questions relates to principles which are similar in nature to (JL), but independent from it. The duty of civility (DC) and justificatory neutrality (JN) once again come to mind (sec. 2.7). Other principles are also interesting candidates, such as

Public Reason Liberalism (PRL). If a law *L* is legitimate, then it is true that there is an argument or reason *R* in favour of *L*, which is publicly given by the government and which each citizen accepts.

(PRL) can be true even if (JL) is not, and vice versa. But while I emphasised that these principles are *logically* independent, the values underlying (DC), (JN) and (PRL) will be similar to the values standing behind (JL). Much work is still to be done to highlight the connections between these principles and their common foundations.

⁷⁰ A recent, sophisticated defense of something like (IT) can be found in Mason 2010. Cf. D’Agostino 1996.

Concluding Remarks

Third, there are many open questions with regard to the various voluntaristic attitudes and their normative significance. I only sketched briefly the variety and richness we can find here (sec. 2.4 & 2.5). Virtually all commentators seem to assume that there is only one relevant attitude – “consent” or “acceptance”. This assumption needs to be abandoned.

Furthermore, I have not provided a deep explanation why expression of the various voluntaristic attitudes is held to be normatively significant. It is an open, and quite difficult, question what explains their significance. Answering it would bring us closer to understanding the attractiveness of justificatory liberalism.

Fourth, I limited myself in this thesis primarily to forms of justificatory liberalism that are based on equal respect, and focus on government acts of *coercion*. But as I hinted at various points (esp. secs. 2.1, 5.3), different lines of argument for (JL) can be imagined, and have recently been sketched (Lister 2008, Bird 2011, McCammon 2011). These positions are still in their infancy, and it is difficult to assess whether they will ultimately turn out to be plausible. But they open up tantalizing and new avenues of inquiry.

To conclude: even though I have argued that (JL) fails, there are many interesting claims and positions in the vicinity of (JL) that are interesting and have high initial appeal. I hope to pursue these questions in future research.

Works Cited

- Archard, David. 2000. [Review of Klosko 2000.] *Philosophy* 75 (294): 622–626.
- Arneson, Richard. 2003. "Liberal Neutrality on the Good: An Autopsy." In *Perfectionism and neutrality: essays in liberal theory*, ed. George Klosko and Steven Wall, 191–218. Lanham, Md.: Rowman & Littlefield.
- . 2011. "Drawing the Baseline." *Cato Unbound*. <http://www.cato-unbound.org/2011/10/19/richard-arneson/drawing-the-baseline/>.
- Barry, Brian. 1995. *Justice as impartiality*. Oxford: Clarendon Press.
- Besson, Samantha. 2005. *The Morality of Conflict: Reasonable Disagreement and the Law*. Oxford: Hart.
- Bird, Colin. 2011. "Coercion and Public Justification." *SSRN eLibrary* (March 28). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1767009.
- Blake, Michael. 2001. "Distributive Justice, State Coercion, and Autonomy." *Philosophy & Public Affairs* 30 (3): 257–296.
- Bohman, James, and Henry Richardson. 2009. "Liberalism, Deliberative Democracy, and 'Reasons That All Can Accept'." *Journal of Political Philosophy* 17 (3): 253–274.
- Brettschneider, Corey. 2007. "Unreasonable Disagreement: Reply to Lovett." *Political Theory* 35 (6): 811–815.
- Brower, Bruce. 1994. "The Limits of Public Reason." *The Journal of Philosophy* 91 (1): 5–26.
- van der Burg, Wibren, and Frans Brom. 1999. "Eine Verteidigung der Staatlichen Neutralität." In *Angewandte Ethik in der Pluralistischen Gesellschaft*, ed. Klaus Peter Rippe, 53–82. Freiburg: Universitätsverlag.
- Buss, Sarah. 2008. "Personal Autonomy." In *The Stanford Encyclopedia of Philosophy*, ed. Edward Zalta. Fall 2008. <http://plato.stanford.edu/archives/fall2008/entries/personal-autonomy/>.
- Caney, Simon. 1991. "Consequentialist Defences of Liberal Neutrality." *The Philosophical Quarterly* 41 (165): 457–477.
- . 1995. "Anti-perfectionism and Rawlsian Liberalism." *Political Studies* 43 (2): 248–264.
- . 1998. "Liberal Legitimacy, Reasonable Disagreement and Justice." *Critical Review of International Social and Political Philosophy* 1 (3): 19–36.
- Chan, Joseph. 2000. "Legitimacy, Unanimity, and Perfectionism." *Philosophy & Public Affairs* 29 (1): 5–42.

Works Cited

- Christiano, Thomas. 1998. [Review of Gaus 1996.] *The Philosophical Review* 107 (3): 455–458.
- . 2008. *The constitution of equality: democratic authority and its limits*. Oxford: Oxford University Press.
- . 2009. "Must Democracy Be Reasonable?" *Canadian Journal of Philosophy* 39 (1): 1–34.
- Christman, John. 1988. "Constructing the Inner Citadel: Recent Work on the Concept of Autonomy." *Ethics* 99 (1): 109–124.
- . 2011. "Autonomy in Moral and Political Philosophy." In *The Stanford Encyclopedia of Philosophy*, ed. Edward Zalta. Spring 2011. <http://plato.stanford.edu/archives/spr2011/entries/autonomy-moral/>.
- Christman, John, and Joel Anderson, eds. 2005. *Autonomy and the challenges to liberalism new essays*. Cambridge: Cambridge University Press.
- Clarke, Simon. 1999. "Contractarianism, Liberal Neutrality, and Epistemology." *Political Studies* 47 (4): 627–642.
- Cohen, G. A. 2003. "Facts and Principles." *Philosophy & Public Affairs* 31 (3): 211–245.
- Colburn, Ben. 2010. "Anti-Perfectionisms and Autonomy." *Analysis* 70 (2): 247–256.
- Copp, David. 2011. "Reasonable Acceptability and Democratic Legitimacy: Estlund's Qualified Acceptability Requirement." *Ethics* 121 (2): 239–269.
- Cummiskey, David. 1990. "Kantian Consequentialism." *Ethics* 100 (3): 586–615.
- D'Agostino, Fred. 1996. *Free Public Reason: Making It up as We Go*. New York: Oxford University Press.
- Darwall, Stephen. 1977. "Two Kinds of Respect." *Ethics* 88 (1): 36–49.
- Dryzek, John, and Simon Niemeyer. 2006. "Reconciling Pluralism and Consensus as Political Ideals." *American Journal of Political Science* 50 (3): 634–649.
- Dworkin, Ronald. 1981. *Taking Rights Seriously*. New Impression with a Reply to Critics. London: Duckworth.
- Eberle, Christopher. 2002. *Religious Conviction in Liberal Politics*. Cambridge: Cambridge University Press.
- Enoch, David. 2009. "On Estlund's Democratic Authority." *Iyyun: The Jerusalem Philosophical Quarterly* 58: 35–48.
- Estlund, David. 1998. "The Insularity of the Reasonable: Why Political Liberalism Must Admit the Truth." *Ethics* 108 (2): 252–275.
- . 1999. [Review of Gaus 1996.] *Philosophy and Phenomenological Research* 59 (3): 821–825.

Works Cited

- . 2008. *Democratic Authority: a Philosophical Framework*. Princeton: Princeton University Press.
- Feinberg, Joel. 1989. "Autonomy." In *The Inner citadel: essays on individual autonomy*, ed. John Philip Christman, 27–53. New York: Oxford University Press.
- Feltham, Brian. 2007. "Unreasonable Rejectability and Permissible Coercion." *Proceedings of the Aristotelian Society* 107: 395–401.
- Forst, Rainer. 2007. *Das Recht auf Rechtfertigung. Elemente einer konstruktivistischen Theorie der Gerechtigkeit*. Frankfurt: Suhrkamp.
- Frankena, William. 1986. "The Ethics of Respect for Persons." *Philosophical Topics* 14 (2): 149–67.
- Freeman, Samuel. 1991. "Contractualism, Moral Motivation, and Practical Reason." *The Journal of Philosophy* 88 (6): 281–303.
- . 2006. "Moral Contractarianism as a Foundation for Interpersonal Morality." In *Contemporary Debates in Moral Theory*, ed. James Dreier, 57–76. Malden: Blackwell.
- . 2007a. "The Burdens of Public Justification: Constructivism, Contractualism, and Publicity." *Politics, Philosophy & Economics* 6 (1): 5–43.
- . 2007b. *Rawls*. London: Routledge.
- Friedman, Marilyn. 2000. "John Rawls and the Political Coercion of Unreasonable People." In *The Idea of a Political Liberalism: Essays on Rawls*, ed. Victoria Davion and Clark Wolf, 16–33. Lanham: Rowman & Littlefield.
- Gaertner, Wulf. 2009. *A Primer in Social Choice Theory*. Rev. ed. Oxford: Oxford University Press.
- Galston, William A. 1991. *Liberal purposes: goods, virtues, and diversity in the liberal state*. Cambridge: Cambridge University Press.
- Gaus, Gerald. 1996. *Justificatory liberalism: an essay on epistemology and political theory*. Oxford: Oxford University Press.
- . 2003. "Liberal Neutrality: A Compelling and Radical Principle." In *Perfectionism and neutrality: essays in liberal theory*, ed. George Klosko and Steven Wall, 137–166. Lanham, Md.: Rowman & Littlefield.
- . 2010a. "Coercion, Ownership, and the Redistributive State: Justificatory Liberalism's Classical Tilt." *Social Philosophy and Policy* 27 (1): 233–275.
- . 2010b. "On Two Critics of Justificatory Liberalism: A Response to Wall and Lister." *Politics, Philosophy & Economics* 9 (2): 177–212.
- . 2011. *The order of public reason: a theory of freedom and morality in a diverse and bounded world*. Cambridge: Cambridge University Press.

Works Cited

- Gaus, Gerald, and Kevin Vallier. 2009. "The Roles of Religious Conviction in a Publicly Justified Polity." *Philosophy & Social Criticism* 35 (1-2): 51–76.
- Goodin, Robert. 2007. "Enfranchising All Affected Interests, and Its Alternatives." *Philosophy & Public Affairs* 35 (1): 40–68.
- Gosseries, Axel. 2010. "Publicity." In *The Stanford Encyclopedia of Philosophy*, ed. Edward Zalta. Fall 2010. <http://plato.stanford.edu/archives/fall2010/entries/publicity/>
- Greenawalt, Kent. 1993. "On Public Reason." *Chicago-Kent Law Review* 69: 669–689.
- Grotefeld, Stefan. 2006. *Religiöse Überzeugungen im liberalen Staat: protestantische Ethik und die Anforderungen öffentlicher Vernunft*. Stuttgart: Kohlhammer.
- Habermas, Jürgen. 1998. *Faktizität und Geltung*. Erweiterte Ausgabe. Frankfurt: Suhrkamp.
- . 2004. *Wahrheit und Rechtfertigung*. Erweiterte Ausgabe. Frankfurt: Suhrkamp.
- Kelly, Erin, and Lionel McPherson. 2001. "On Tolerating the Unreasonable." *Journal of Political Philosophy* 9 (1): 38–55.
- Klosko, George. 1997. "Political Constructivism in Rawls's Political Liberalism." *The American Political Science Review* 91 (3): 635–646.
- . 2000. *Democratic procedures and liberal consensus*. Oxford; New York: Oxford University Press.
- . 2003. "Reasonable Rejection and Neutrality of Justification." In *Perfectionism and neutrality: essays in liberal theory*, ed. George Klosko and Steven Wall, 167–190. Lanham, Md.: Rowman & Littlefield.
- Larmore, Charles. 1990. "Political Liberalism." *Political Theory* 18 (3): 339–360.
- . 1994. "Pluralism and Reasonable Disagreement." *Social Philosophy and Policy* 11 (1): 61–79.
- . 1999. "The Moral Basis of Political Liberalism." *The Journal of Philosophy* 96 (12): 599–625.
- Lecce, Steven. 2003. "Contractualism and Liberal Neutrality: A Defence." *Political Studies* 51 (3): 524–541.
- . 2008. *Against perfectionism: defending liberal neutrality*. Toronto: University of Toronto Press.
- Lijphart, Arend. 1999. *Patterns of democracy: government forms and performance in thirty-six countries*. New Haven: Yale University Press.
- List, Christian. 2002. "Two Concepts of Agreement." *The Good Society* 11 (1): 72–79.
- . 2011. "The Logical Space of Democracy." *Philosophy & Public Affairs* 39 (3): 262–297.

Works Cited

- Lister, Andrew. 2007. "Public Reason and Moral Compromise." *Canadian Journal of Philosophy* 37 (1): 1–34.
- . 2008. "Public Reason and Democracy." *Critical Review of International Social and Political Philosophy* 11 (3): 273–289.
- . 2010a. "Liberal Foundations of Democratic Authority." *Representation* 46 (1): 19–34.
- . 2010b. "Public Justification and the Limits of State Action." *Politics, Philosophy & Economics* 9 (2): 151–175.
- Lott, Micah. 2006. "Restraint on Reasons and Reasons for Restraint: A Problem for Rawls' Ideal of Public Reason." *Pacific Philosophical Quarterly* 87 (1): 75–95.
- MacIntyre, Alasdair. 1988. *Whose justice? Which rationality?* Notre Dame: University of Notre Dame Press.
- Manin, Bernard. 1987. "On Legitimacy and Political Deliberation." *Political Theory* 15 (3): 338–368.
- Mason, Andrew. 1993. *Explaining political disagreement*. Cambridge: Cambridge University Press.
- . 2010. "Rawlsian Theory and the Circumstances of Politics." *Political Theory* 38 (5): 658–683.
- May, Simon Căbulea. 2005. "Principled Compromise and the Abortion Controversy." *Philosophy & Public Affairs* 33 (4): 317–348.
- . 2009. "Religious Democracy and the Liberal Principle of Legitimacy." *Philosophy & Public Affairs* 37 (2): 136–170.
- McCabe, David. 2000. "Knowing About the Good: A Problem with Antiperfectionism." *Ethics* 110 (2): 311–338.
- McCammon, Christopher. 2011. "Republican Foundations for Liberal Restraint." <http://www.philosophy.northwestern.edu/conferences/moralpolitical/11/papers/McCammon.pdf>.
- Moore, Margaret. 1996. "On Reasonableness." *Journal of Applied Philosophy* 13 (2): 167–178.
- Morris, Christopher. 2012. "State Coercion and Force." *Social Philosophy and Policy* 29 (1): 28–49.
- Nagel, Thomas. 1987. "Moral Conflict and Political Legitimacy." *Philosophy & Public Affairs* 16 (3): 215–240.
- . 1991. *Equality and partiality*. New York: Oxford University Press.
- . 2005. "The Problem of Global Justice." *Philosophy & Public Affairs* 33 (2): 113–147.
- Neumann, Michael. 2004. "Can't We All Just Respect One Another a Little Less?" *Canadian Journal of Philosophy* 34 (4): 463–484.
- Nickel, James. 1982. "Equal Respect and Human Rights." *Human Rights Quarterly* 4 (1): 76–93.
- Niemeyer, Simon, and John Dryzek. 2007. "The Ends of Deliberation: Meta-consensus and Inter-subjective Rationality as

Works Cited

- Ideal Outcomes." *Swiss Political Science Review* 13 (4): 497–526.
- Nussbaum, Martha. 2011. "Rawls's Political Liberalism. A Reassessment." *Ratio Juris* 24 (1): 1–24.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- . 2011. *On What Matters*. Vol. 1. Oxford: Oxford University Press.
- Pettit, Philip. 1989. "Consequentialism and Respect for Persons." *Ethics* 100 (1): 116–126.
- Pogge, Thomas. 2001. "What We Can Reasonably Reject." *Philosophical Issues* 11: 118–147.
- Price, Terry. 2000. "Epistemological Restraint—Revisited." *Journal of Political Philosophy* 8 (3): 401–407.
- Quong, Jonathan. 2005. "Disagreement, Asymmetry, and Liberal Legitimacy." *Politics, Philosophy & Economics* 4 (3): 301–330.
- . 2011. *Liberalism Without Perfection*. Oxford: Oxford University Press.
- Rawls, John. 1999a. "The Idea of Public Reason Revisited." In *Collected Papers*, ed. Samuel Freeman. Cambridge and London: Harvard University Press.
- . 1999b. *A Theory of Justice*. Revised edition. Cambridge: Harvard University Press.
- . 2001. *Justice as fairness: a restatement*. Cambridge: Harvard University Press.
- . 2005. *Political Liberalism*. Expanded Edition. New York: Columbia University Press.
- Raz, Joseph. 1988. *The Morality of Freedom*. New Edition. Clarendon Press.
- . 1990. "Facing Diversity: The Case of Epistemic Abstinence." *Philosophy & Public Affairs* 19 (1): 3–46.
- . 1998. "Disagreement in Politics." *American Journal of Jurisprudence* 43: 25–52.
- . 2001. *Value, respect, and attachment*. Cambridge: Cambridge University Press.
- Reidy, David. 2007. "Reciprocity and Reasonable Disagreement: From Liberal to Democratic Legitimacy." *Philosophical Studies* 132 (2): 243–291.
- Rescher, Nicholas. 1993. *Pluralism: against the demand for consensus*. Oxford: Oxford University Press.
- Ripstein, Arthur. 2004. "Authority and Coercion." *Philosophy & Public Affairs* 32 (1): 2–35.
- Sandel, Michael. 1994. [Review of Rawls, John. 1993. *Political Liberalism*.] *Harvard Law Review* 107 (7): 1765–1794.
- . 1998. *Liberalism and the Limits of Justice*. Second ed. Cambridge: Cambridge University Press.

Works Cited

- Saunders, Ben. 2010. "Democracy, Political Equality, and Majority Rule." *Ethics* 121 (1): 148–177.
- Scanlon, Thomas. 1982. "Contractualism and Utilitarianism." In *Utilitarianism and Beyond*, ed. Amartya Sen and Bernard Williams. Cambridge: Cambridge University Press.
- . 1998. *What We Owe To Each Other*. Cambridge and London: Harvard University Press.
- . 2011. "How I Am Not A Kantian." In *On What Matters*, by Derek Parfit, Vol. 2, 116–142. Oxford: Oxford University Press.
- Schwartzman, Micah. 2004. "The Completeness of Public Reason." *Politics, Philosophy & Economics* 3 (2): 191–220.
- Simmons, John. 2009. "Philosophical Anarchism." *SSRN eLibrary* (February 16). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1344425.
- Solum, Lawrence. 1993. "Constructing an Ideal of Public Reason." *San Diego Law Review* 30: 729–762.
- Southwood, Nicholas. 2010. *Contractualism and the Foundations of Morality*. Oxford: Oxford University Press.
- Suikkanen, Jussi. 2005. "Contractualist Replies to the Redundancy Objections." *Theoria* 71 (1): 38–58.
- Vallier, Kevin. 2011. "Liberal Politics and Public Faith: A Philosophical Reconciliation". Ph.D. thesis, University of Arizona. http://www.kevinvallier.com/wp-content/uploads/2011/04/Kevin_Vallier_DissertationElectronic.pdf.
- Waldron, Jeremy. 1987. "Theoretical Foundations of Liberalism." *The Philosophical Quarterly* 37 (147): 127–150.
- . 1993. "Religious Contributions in Public Deliberation." *San Diego Law Review* 30: 817–848.
- . 1994. "Disagreements About Justice." *Pacific Philosophical Quarterly* 75: 372–387.
- . 1999. *Law and Disagreement*. Oxford: Clarendon Press.
- . 2005. "Moral Autonomy and Personal Autonomy." In *Autonomy and the challenges to liberalism: new essays*, ed. John Philip Christman and Joel Anderson. Cambridge: Cambridge University Press.
- . 2006. "The Core of the Case Against Judicial Review." *Yale Law Journal* 115: 1346–1406.
- Wall, Steven. 1996. *Liberalism, Perfectionism and Restraint*. University of Oxford.
- . 2002. "Is Public Justification Self-Defeating?" *American Philosophical Quarterly* 39 (4): 385–394.
- . 2010a. "On Justificatory Liberalism." *Politics, Philosophy & Economics* 9 (2): 123–149.

Works Cited

- . 2010b. "Neutrality for Perfectionists: The Case of Restricted State Neutrality." *Ethics* 120 (2): 232–256.
- Weithman, Paul. 2002. [Review of Klosko 2000.] *Ethics* 112 (3): 621–626.
- Wenar, Leif. 1995. "Political Liberalism: An Internal Critique." *Ethics* 106 (1): 32–62.
- Westmoreland, Robert. 1999. "The Truth About Public Reason." *Law and Philosophy* 18 (3): 271–296.
- Williams, Andrew. 1998. "Incentives, Inequality, and Publicity." *Philosophy & Public Affairs* 27 (3): 225–247.
- . 2000. "The Alleged Incompleteness of Public Reason." *Res Publica* 6 (2): 199–211.
- Wolterstorff, Nicholas. 1997. "The Role of Religion in Decision and Discussion of Political Issues." In *Religion in the public square: the place of religious convictions in political debate*, ed. Robert Audi and Nicholas Wolterstorff, 67–120. Lanham, Md.: Rowman & Littlefield Publishers.
- Wood, Allen. 1999. *Kant's Ethical Thought*. Cambridge: Cambridge University Press.