

## Chapter 5. The ethics of implementation

*[[WORK IN PROGRESS – DO NOT CIRCULATE]]*

*Jojanneke Vanderveen, Vrije Universiteit Amsterdam, 1 September 2017*

*j.s.vanderveen@vu.nl*

In chapter 3, I presented a distinction between necessary and circumstantial requirements and maintained that the latter are the appropriate requirements for judgments of political legitimacy. I gave a conceptual argument for circumstantial requirements. Now, I want to present a normative argument for circumstantial requirements. The normative argument makes it more urgent to abandon necessary requirements in favor of circumstantial requirements. I present a number of values that power holders should observe when asking themselves which uses of coercive power can be morally justified. These values derive from the purpose of the state.

### *5.1. Where values of implementation come in*

Briefly put, the main normative problem for necessary requirements is that following up on them may have bad consequences. Democracy may be required by political justice, but if, for example, implementing the institutions of a democracy creates large-scale chaos and can be expected to cost many lives, then there is something to say for not protecting justice ‘though the heavens fall’ (Geuss 2008, 83, Nielsen 2007, 22).<sup>1</sup> As G.A. Cohen says: “I don’t see how anyone, whatever she thinks justice is, can deny the possibility that certain facts, or other values, might make it inappropriate, or too difficult, or too costly, to produce justice” (Cohen 2008, 302); “[j]ust as truth is not a necessary condition of all justifiable utterance, so it is sometimes justifiable, all things considered, to deviate from justice in the formation of social institutions” (304). The same holds for the maintenance of and changes in institutions, and for the implementation of policies. Policies, like institutions,

---

<sup>1</sup> As Michael Walzer says in his influential article “Political Action: The Problem of Dirty Hands”: “we would not want to be governed by men who consistently adopt that [‘absolutist’, non-rights-violating] position” (Walzer 1973, 162).

may be thought desirable or required on the basis of first-order considerations (e.g.: in its party program, a party may propose to increase spending on foreign aid substantially due to concerns of distributive justice), but undesirable when it comes to implementation (e.g.: the only way to spend the money may be to hand it to corrupt regimes). **[[ADD W.D. ROSS]]**

If Cohen is correct in saying that there may be reasons to deviate from justice in making decisions about implementation – which I submit he is – then we should look for the normative considerations that apply to the implementation of institutions and policies that do not yet come into view when we decide whether we find them desirable to begin with, but that do enter the scene when we are ready to make actual decisions. (Throughout, we must reason from the perspective of rulers. When I say that ‘we’ make decisions, we must imagine ourselves to have a position of sufficient power to actually make these decisions.) If we find an institution worth striving for, but if the fact that it is worth striving for is not enough to determine whether it would be morally desirable to implement it *now*, then there must be reasons – *moral* reasons – that explain this gap. The question is thus: which moral considerations apply to this gap between determining which goals we wish to attain and what we should implement now? I will refer to these considerations as the values of implementation. These are the values that enter at the all-things-considered level of moral judgment.

In the previous chapter, we already recognized this gap, but for non-normative reasons; it may be factually *impossible* (for the time being) to implement something that is worth striving for. Call this the problem of technical feasibility. Now, we discuss the normative content of the gap. Recognizing that normative considerations apply to the gap between plans and implementation opens up the possibility that a proposal that has moral reasons in favor of it, ought not to be implemented, because its implementation is morally undesirable under current circumstances. The *prima facie* reasons that count in favor of an institution may be outweighed by the negative moral value that it creates if it would be implemented now. Call this the problem of normative feasibility; it may not be feasible to obtain the

moral value of a *prima facie* morally desirable institution.<sup>2</sup> Of course, it is not thereby said *what* might make implementation of a (first-order) desirable plan (second-order) undesirable. In the extreme case, one might argue that political leaders should not let anything stand in their way regarding the implementation of things that have been determined to be desirable as a goal, on e.g. the adage that “*fiat justitia, pereat mundus*”. If we find such an adage implausible and think it is not legitimate for rulers to close their eyes for the moral costs of their decisions (cf. Galston 2010, 392), then we apparently use additional moral principles – principles of implementation.

These principles govern what happens *while* we pursue whatever goals we have. The relevance of these principles stems from the fact that the end does not justify the means. At the very least, the end cannot provide a full justification for the use of coercive power, and there are reasons to care what happens *while* we pursue our ends. Here, we are specifically interested in the values that agents of the state, those that are in charge of exercising its coercive power, need to take into account while they pursue their aims.

It may be helpful, here, to distinguish between three different categories within normative theorizing. The following passage by Allan Buchanan sets these out.

The task of ideal theory is to set the most important and most distant moral targets for a better future, the ultimate standards for evaluating current international law. Nonideal theory’s task is to guide our efforts to approach those ultimate targets, both by setting intermediate moral targets, as way-stations on the path toward the ultimate standards laid down by ideal theory, and by helping us to determine which means and processes for achieving them are morally permissible. (Buchanan 2004, 60-1)

In Buchanan’s characterization, ultimate targets and intermediate targets belong, respectively, to ideal and non-ideal theory. Given that both these types of normative theorizing produce evaluative tools, targets to aim for, neither of these provides an all-thing-considered perspective. The latter category – morally permissible means and processes – concerns the level of implementation, and hence belongs to the level we are interested in.

---

<sup>2</sup> For the distinction between technical and normative feasibility, see e.g. Erez (2015, 47-8).

My claims in this chapter should be understood to be relatively modest. My aim is not to say that the values of implementation I point to *do* provide a full justification for the use of coercive power, and that the moral value of the end that is pursued is of no import. I do not address the way in which the value of the end and the impact of the means interplay. Rather, I maintain that it is important to consider the question of how power may be used independently of the end towards which it is used. We should distinguish between the desirability of the end and the desirability of the means (Räikkä 1998, 33). I am interested in what it is that makes means desirable or undesirable. Giving an account of these values will aid us in making ‘route comparisons’ (Carey 2015, 154). As Valentini points out, this question is still “vastly under-explored” in political theory (Valentini 2012b, 662).

Of course, the values that inform our ends likely also influence our judgments about means. If we strive for a free society, we will probably take issue with severe curtailments of freedom on the way to this free society – even if such curtailments would really get us to a state of freedom quickest (which may be implausible, but that is another issue). The point is that formulating a goal is not the same as formulating the moral values and principles that govern morally justified movement towards a goal. It is a different question. This question brings its own moral challenges, especially if coercive power is used. Acknowledging these challenges does not undermine the relevance of formulating worthy goals. Formulating goals is also a relevant challenge to take on. It is simply that the question of implementation requires separate attention. When we speak about the “moral costs of changeover” (Räikkä 1998, 33, Buchanan 2002, 62, Gilabert 2012, 52, Erez 2015, 49), what kind of costs do we have in mind? By reference to which values can we get a handle on what counts as permissible, and what does not?

Rawls’ ‘transitional theory’ provides, to some extent, a conceptual tool to describe what we are doing. His concept of transitional theory concerns any transition from a society that is not well-ordered to one that is, starting from unfavorable conditions (Rawls 1999a, 90). Rawlsian transitional theory consists, then, of two parts of which we only look at one. Where Rawls thinks of transitional theory as (1) a transition to (2) well-orderedness, we leave aside

for now that second part, and ask only whether something can be said about transitioning as a concern of its own. We do, then, work on the assumption that this is an intelligible and useful question. I thereby reject Rawls' claim that, without having an ideal in mind, "nonideal theory [of which transitional theory is a part] lacks an objective, an aim, by reference to which its queries can be answered" (Rawls 1999a, 90). There are questions of transition that do not pertain to the goal one is working towards, but to the road that one walks while trying to realize the goal. Are there roads we should not walk? How do we judge this?

In summary: thinking about the values of implementation is thinking about the values we use to judge the roads political leaders make our societies walk, and the means they use to build these roads. While the adequacy of a road of course partly depends on where it aims to go – e.g. a state that meets Rawls' principles of justice, or Valentini's principles of democracy – there is reason to think the road in itself can be better or worse, and that the quality of the road is of moral significance. But what kind of values can we use to judge this quality? That is our question.

Before we move on to our discussion of the values of implementation, we should ask ourselves how come something that is a good evaluative standard nevertheless does not prescribe what we should implement now. This will put the issue into a clearer perspective.

### *5.2. Causes for the gap*

The gap between evaluative and prescriptive norms may be larger or smaller, depending on how utopian one's evaluative thinking is. The gap between evaluative and prescriptive thinking increases as one's evaluative ideals diverge more from the status quo. A utopian example will hence most clearly illustrate what may cause the gap.

Take, e.g., Nozick's framework for utopias. In Nozick's ideal, the world would be "a place where people are at liberty to join together voluntarily to pursue and attempt to realize their own vision of the good life in the ideal community where no one can *impose* his own utopian vision upon others"

(Nozick 1974, 312). If we adopt Nozick's vision, we favor institutions that are radically different from the ones that currently exist anywhere. The implementation of our proposal would require the abolishment of the state as we know it and the formation of an infrastructure that allows people to rearrange into like-minded groups, and set up their own institutions.

When we assess how appealing we find this vision, we are engaging in first-order reasoning; we are assessing the merits of such a world, based for example on our evaluation of the importance of freedom and our expectations about the extent to which humans might be capable of adapting to functioning under such a structure. If we value freedom greatly, and expect that it is in principle possible for humans to function well in this set-up, we might adopt the ideal.

Now imagine that we enter the political arena (say we are the Voluntarist Party) with this ideal, and manage to obtain a position of power. Being wholly politically inexperienced, we now intend to abolish all institutions, to replace them with just one agency that deals with complaints about coercion (and is granted the authority to coerce coercers not to coerce). The following things might happen, which might just make this (gigantic) move morally objectionable.

#### 5.2.1. The circumstances are not right.

It may well be that the circumstances for implementation of our framework are not right. In the previous chapter, we considered circumstances that make it *impossible* to implement some arrangement (democracy, in that case). Even if it is, technically speaking, *possible*, however, the circumstances may be uncongenial. Such a radical change might, for example, upset relationships with other countries (because we overthrow trade agreements, for example), or hamper ongoing projects within our country that are considered valuable (e.g. infrastructural projects) due to at least temporary lack of clarity of responsibility. This lack of readiness of the circumstances may cause problems even if everybody in the country agrees that the framework is desirable, evaluatively speaking. The problems are *moral* (and not just practical) problems if they have a morally relevant impact, such as when people's well-

being or their ability to shape their own life is undermined. The not-rightness of circumstances will often be characterized by the chaos that follows the implementation of an arrangement. Chaos is likely to be more severe if the change is more radical. The not-rightness of circumstances can be regarded as a form of normative infeasibility, related to the moral costs of transitioning (Räikkä 1998, 38, Erez 2015, 48-9).

### 5.2.2. There is disagreement.

Chaos can arise even if everybody stands behind the plan; it can easily result from simple lack of coordination and clarity about responsibilities. When it comes to political issues, however, there will usually be *disagreement*. (And the more utopian a proposal is, the more opposition it will likely face.)

In political philosophy, disagreement is widely recognized as a normatively relevant fact. Some theorists argue that disagreement is a constitutive assumption of the whole field of normative political theory, and that political philosophy cannot be conducted independently of acknowledgement of the existence of disagreement (Williams 2005, 77 defends this position, cf. also List and Valentini 2016, 546-7, Larmore 2013, 278). For Rawls, the fact that people disagree over comprehensive doctrines is the reason to recast (in *Political Liberalism* (1993, xvii)) his theory of justice as fairness (as presented in *A Theory of Justice*) as a theory that is ‘political, not metaphysical’ (Rawls 1985). That is to say: Rawls later thought that a conception of justice should be acceptable to all those who endorse reasonable comprehensive doctrines, but who nevertheless disagree on metaphysical issues. It is fair to say that dealing with disagreement has become a major issue in contemporary political theory (e.g. Caney 1998, Reidy 2007, Valentini 2013, Gaus 2011, ...).

How the issue of disagreement is conceived and how it is incorporated into theory, however, allows for many different possibilities. The question for now is: what moral considerations does disagreement raise when it comes to implementation, and has the issue of disagreement in the literature been addressed with an eye to this question? To see what our task is when proposing values of implementation, and in how far the existing literature is helpful, we

should dwell on these questions, which is what I do in the remainder of this section.

Following Rawls, many liberal theorists have, over the last decades, focused on the issue of disagreement on moral questions, and the impact such disagreement should have on political institutions. This interest is motivated by the worry that, if people disagree on what morality requires, and if this would induce them to make different and opposing political decisions, there is a problem in justifying the prevalence of one view over other views. Why, these theorists ask, if we start from liberal values, should one moral doctrine have the right to impose its views on others? In response to this question, roughly two responses dominate the field. They are the responses of the theories we have discussed in the previous chapter. I will now discuss them with an eye to the problem of disagreement.

- Avoidance

The Rawlsian answer to the question of what to do in light of disagreement is: avoid those values and principles that are contentious – the method of avoidance (Rawls (1985, 231), Young (2006, 163), Caney (1998, 20), Hampton (1989, 799)). What is left after contentious values and principles have been avoided, he calls the ‘overlapping consensus’. This overlapping consensus contains political values on which all comprehensive ethical views – i.e.: ethical views that provide regulative principles for a complete moral life – which are reasonable, supposedly (would) agree (1993, 147). He expects that the principles that end up in this overlapping consensus can help reach agreement on constitutional essentials and the basic questions of justice (156). A similar theory is offered by Charles Larmore who, instead of referring to ‘the fact of reasonable pluralism’ as Rawls does, speaks of “reasonable disagreement”, and hopes that those who reasonably disagree on what makes life worth living, may still agree on a core of morality that can attract the assent of reasonable people (Larmore 1996, 12-13).

The strategy is hence to exclude disagreement from politics; what we disagree on, is left to people’s own discretion. What we agree on is what we settle collectively. As Rawls says in *Political Liberalism*: “the knowledge and

ways of reasoning that ground our affirming the principles of justice and their application to constitutional essentials and basic justice are to rest on *the plain truths now widely accepted, or available, to citizens generally* – theories that are “in dispute” are avoided (Rawls 1993, 225, my emphasis). Thus, citizens refrain from forcing their own doctrines on others insofar as these are controversial.

Such a theory, even if it is consensus-oriented in what it calls the political sphere, can be seen as accepting (descriptive) pluralism because the consensus exists among people of differing (reasonable) philosophies and religions. There are likely to be many values and principles on which they could or do come into conflict, but on these matters they agree to tolerate each other and they do not make political issues out of these. In this sense, society remains pluralist despite a politics of consensus.

Theories like Rawls’ and Larmore’s rely on the assumption that something of import will be captured by the overlapping consensus or the core morality; if there were nothing important that could count on consensus and was uncontroversial, then the politics of overlapping consensus would be doomed. We should note, however, that the consensus Rawls and Larmore envision is not unqualified. They both stress the importance of reaching *reasonable* consensus. Hence, for theories like Rawls’ and Larmore’s, a lot hinges on what this standard of reasonableness is taken to be. Rawls’ overlapping consensus is thought to hold among *reasonable* doctrines. Those who hold reasonable doctrines will, by the use of public reason (Rawls 1993, 137), be able to endorse certain ways of the exercise of power concerning constitutional essentials and basic questions of justice, which are captured by the institutions agreed upon in the overlapping consensus (Rawls 1993, 140). The legitimacy of coercive power depends on the content of the overlapping consensus, which is established by what reasonable people can agree on. Its content must not be bent to ‘existing unreason’ (Rawls 1993, 144). It is necessary, then, that those who are regarded as reasonable by Rawls, be able to accept the regulative principles by which power is exercised. It is not necessary for political legitimacy that those whom Rawls regards as unreasonable are able to accept these principles (Reidy 2007, 260).

The weight is thus on the concept of reasonableness; “[i]f political legitimacy is equal acceptability by reasonable people, its substantive profile depends on, and varies with, the content of the idea of reasonableness that we suppose” (Besch 2013, 62-3). Regarding the issue of the position of the unreasonable, Rawls does not say much, except in a footnote in which he argues that we have the task of “containing [unreasonable doctrines] – like war and disease – so that they do not overturn political justice” (1993, 64). I will return to the issue of reasonable versus unreasonable disagreement.

The method of avoidance responds to disagreement by restricting the political sphere; only those values and principles that can count on reasonable consensus are addressed publicly. These approaches hence aim at eliminating disagreement (from the political sphere), rather than adjudicating it. They look for tenets of social justice that can acquire everybody’s assent. Other theorists take another approach, and instead propose procedures that settle disagreements one way or another, rather than avoiding them. They seek for the right response to disagreement in *political* justice.

- Procedures

If we do not want to restrict the political agenda to values and principles that can count on (reasonable) agreement, we may instead look for procedures to make satisfactory decisions given that people do not agree what this decision should be. This is what theorists who focus on political justice as the right response to disagreement do. The thought is that while we may not agree what the right decision is, we may still agree on a procedure to adjudicate our differences. Such a procedure would be politically just, according to e.g. Valentini and Pettit (Valentini 2012a, 600, Pettit 2015, 13). [[Check: Pettit says the procedure needs to attract democratic approval, while Valentini says the procedure needs to be democratic. This is not the same; procedures that everybody agrees to need not be democratic, while democratic procedures need not attract everybody’s support.]]

In contrast to theorists who adopt the method of avoidance, such proceduralist accounts do not aim to dispel moral disagreement from the political sphere. Rather, they adopt an overarching way to manage these

disagreements. In this way, they are unlike avoidance approaches. In another way they are alike, however. Like avoidance approaches, the cited procedural approaches limit themselves to responding to *reasonable* disagreement. When we look for adjudication of disagreement, it is disagreements between *reasonable* positions that are intended. Valentini, for example, discounts “unreasonable disagreement in so far as this falls outside the liberal commitment to equal respect. Liberals are committed to equal respect *qua* mutual justification, but their justificatory audience does not include those who, from a liberal perspective, hold unreasonable views. [...] Not every objection carries normative force, only reasonable ones do” (Valentini 2013, 186). The problem they address is how the imposition of coercive laws on those who do not agree with them may be justified. In giving this justification, the argument goes, we do not have to try to appeal to those who occupy unreasonable positions; a justification that is acceptable to reasonable people is enough.

Here too, then, the notion of reasonableness plays a central role in reaching normative conclusions. If our purpose here were to assess the plausibility of the conclusions of both avoidance and proceduralist views, we should turn to the question what this notion of reasonableness is, and how its content is determined. This is not our purpose, however. Rather, given that we are interested in the values that are relevant to the ethics of implementation – the ethics of the actual *use* of coercive power – we should ask how far acknowledging reasonable disagreement and proposing a way to deal with it get us towards answering the question which values and normative considerations apply to implementation. Is reasonable disagreement the only type of disagreement that causes a gap between evaluative and prescriptive tools? Our conclusion should be that it is not; unreasonable disagreement matters too.

Despite the brief remarks that for example Rawls and Valentini make about the presence of the unreasonable in society, we would not be reading their theories most charitably if we presumed that they think this issue provided no moral questions when it comes to implementation. Rather, the best way to interpret what they are doing is to understand them as working on

the premise that disagreement is an ineliminable aspect of society, and that even ideal theories of justice – evaluative standards – should hence take at least *reasonable* disagreement into account. On this interpretation of the cited theorists, they would hold that “disagreement in political theory may be viewed, not merely as “epistemic”, reflecting different beliefs about the same truth, but as partly “constitutive” of the correctness conditions [...] of normative claims themselves” (List and Valentini 2016, 547). In other words: whether the conclusions of a normative theory should be accepted, may depend on whether they can be applied under circumstances of reasonable disagreement. I gather, however, that we are here still talking about the normative conclusions of *evaluative* theories. The question is: is acknowledging reasonable disagreement enough to get us to a *prescriptive* theory?

The reasons to regard disagreement as constitutive of the correctness conditions in political theory have to do with making the conclusions of theories acceptable in actual societies. Christian List and Laura Valentini suggest that theorists who take a constitutive rather than an epistemic approach to disagreement (and I interpret Valentini as being one of those theorists) may do so for mainly two reasons. First, to find political principles that may permissibly be enforced, one may think that they need to be acceptable to reasonable people. Second, if a political theory appeals to very few people, its imposition would not be durable and stable, and hence not very practicable if one wants to avoid dictatorial regimes (List and Valentini 2016, 547-8).

Both of these reasons spring from a concern to make political theory relevant for the real world. Now, it should be noted that accepting these reasons does not mean one thinks *sufficient* conditions have been met to make implementation of one’s conclusions morally acceptable. Acceptability to reasonable people may be a necessary, but not a sufficient condition for permissible enforcement. Similarly, such acceptability may be necessary for stability and durability, but one need not think that it is the only factor that determines the success of a political regime. Although looking for principles that may permissibly be enforced may make it seem as though one thinks that

the principles one propounds in the theory may in fact be enforced, we could leave open the possibility that what these theorists claim is that reasonable acceptability is necessary, but maybe not sufficient, for the permissibility of enforcing the precepts of the theory.

If we adopt the necessary-not-sufficient interpretation, then we could take these theories to be claiming that either the proposed overlapping consensus or the proposed procedures should be used as evaluative standards; something is amiss in a society that does not meet this standard. We would, on the other hand, not interpret them as prescribing the implementation of their proposals, given that additional moral considerations may have to be brought to bear on the question of implementation. One of the most important issues that, next to uncongenial circumstances, has to be taken into account in the ethics of implementation is the presence of *unreasonable* disagreement.

The claim that I want to lay most stress on is that unreasonable disagreement matters for political legitimacy. In the remainder of this chapter I argue that political power must be used to fulfill the purpose of the state, and that the purpose of the state includes the interests of all citizens, whether reasonable or not. Moreover, the disagreement of unreasonable citizens may affect the moral consequences of the actions of the state, and hamper the fulfillment of its purpose, even if all reasonable citizens endorse these actions. So what is the purpose of the state?

### 5.3.2. The purpose of the state

Although the question is straightforward – what is the purpose of the state? – answering it is a challenging task. I do not aim to give a comprehensive and definitive answer, but rather to provide reference points that can bring us further in judging the choices of power holders. Let us take a first lead from Rawls. According to Rawls, society is to be a “cooperative venture for mutual advantage” (Rawls 1999b, 4). A society that shares an institutional structure has “the purpose of advancing the good of those taking part in it” (Timmerman 2014, 5).

This line of thought does not start with Rawls. Both Hobbes and Locke, although different in many respects, see the state as serving people’s mutual

advantage. According to Hobbes, the state is intended for people's "*Peace and Common Defence*" (Hobbes 1985, 228 [88]) and is hence in the self-interest of those subject to state authority. On a Hobbesian view, the state constitutes a strategic relationship between its citizens (Christman 2005, 341).

Taking a Lockean line, the state is intended to secure the non-violation of people's pre-institutional rights. When asking why someone would part with their natural freedom, Locke thinks "it is obvious to answer, that though in the state of nature he hath such a right [to natural freedom], yet the enjoyment of it is very uncertain, and constantly exposed to the invasion of others" (§123). It is this uncertainty that

makes him willing to quit a condition, which, however free, is full of fears and continual dangers: and it is not without reason, that he seeks out [...] others [...] for the mutual preservation of their lives, liberties and estates, which I call by the general name, property.

The great and *chief end*, therefore, of men's uniting into common-wealths, and putting themselves under government, *is the preservation of their property.* (§123-4)

Locke's rationale for having a state, then, is similar to Hobbes': it is putting an end to a life full of fear and danger.

Unlike Hobbes, Locke refers to pre-institutional (natural) rights in his defense of the state. In the liberal tradition, many theorists give an account of people's rights and expect the state to protect these rights. Libertarians perhaps stay closest to Locke in the content of the set of pre-institutional rights they recognize (e.g. Nozick 1974, ix, Mack 1990, 519, Otsuka 1998, 66, Gibbard 1976, 77). The human-rights tradition, as exemplified by e.g. Allen Buchanan (2004, chapter 3), also puts focus on a set of pre-institutional moral rights that it is the job of institutions to protect. Human-rights theorists have a broader conception of such rights than libertarians have. They include, for example, positive rights such as the right to subsistence (e.g. Buchanan 2004, 129, Sen 2004, 316, Reidy 2007, 274, Stilz 2012, 255).

Despite the wide variety of views of what exactly constitutes the set of interests that the state ought to protect and promote, a number of values concern all theorists. And although my project here should not be understood as aimed at finding consensus among all accounts, it is nevertheless helpful to

reflect on common denominators. We are looking for values that can guide decisions about the use of coercive power. These values should hence be sufficiently substantive that they can be of use in power holders' deliberations, and sufficiently general that they are relevant over a wide range of decision-making contexts.

Before we move on to set out these values, it is important to note that on any view that explains the rationale for the state in terms of some mutual advantage that participants in the state will gain, the state has an instrumental role. We value the state *provided that* it can deliver this mutual advantage. This means that the primary way to judge how well a state is doing, is to check if it manages to deliver these advantages. Of course, we may put constraints on *how* these advantages may be delivered, but a state that did not even try to provide the relevant advantages would not serve the purpose for which it is valued and could hence rightly be objected to. We justify the use of coercive power by the state, despite a presumption against coercion, by pointing to the advantages such a scheme of power has. If the agents of the state do not act to deliver these advantages, the justification for its use of coercive power is undermined.

If this is right, the role morality of the state is importantly consequentialist in nature. In their employment of the coercive power that the state holds, agents of the state must continuously ask themselves whether they promote the goals from which the state gets its justification. That is: whether they serve the values that it is the purpose of the state to serve. Now to the important question: what are these values?

- Security and cooperation

Rawls sees society as a cooperative venture for mutual advantage. Giving this thought a bit more flesh on the bones by referring to Hobbes and Locke, we saw that the mutual advantage can consist in safety from attacks by others. Plausibly, it can also exist in safety from hardship by other causes than attacks by others, such as all kinds of bad luck. Whether we fear hardship through others' behavior or through other causes, the most important advantage we expect from a state is that it reduces our risk of dying early or of attracting

injury. This is, possibly, the most important task of the state, and one that cannot easily be outweighed by other objectives.

Of course, preventing death and injury need not be the end of the state's task. We may extend the task of the state to creating opportunities for cooperation and for providing security in our ventures. Although it is plausible to accept limits to the justified use of coercive power in fulfillment of such a task, it certainly makes sense to see these more extensive endeavors as part of the state's task. Imagine, then, that state agents are wondering: 'may we use coercive power to create these opportunities for cooperation?' or: 'may we use coercive power to provide people with security in this venture?' The fact that they are aiming at creating opportunities for cooperation or to provide security is a first indication that they are using coercive power in the right ways.

This is by no means a definitive conclusion. Murderers can cooperate and thieves may want security in their ventures. Especially those ventures that threaten the cooperation of others or their security, however, are suspicious. Here, the liberal commitment to equal respect (REFS) becomes important. Because the value of the state consists in its being for the *mutual* advantage of each participant in the state, the state should aim to create circumstances that support everyone's security and cooperative endeavors.

The state may be a threat both to people's security and to their cooperative endeavors. Warding off this threat is probably an important concern for liberal theorists, which explains why emphasis has been put on the *limits* of what the state may do. Setting out what the state may *not* do, is a negative task. But we would not care to have a state at all, if there were not something we could say, *positively*, about what we value it for. Protecting us from malice and misfortune and protecting our cooperative ventures states this purpose in a general and plausible way.

If the above is correct, it follows that the state cannot, when reflecting on how to use power in a justified way, ignore the presence of unreasonable disagreement. Even if we could plausibly maintain that the disagreement of those who hold unreasonable doctrines is not normatively relevant in principle (which I am inclined to doubt), it certainly is normatively relevant

when it comes to consequences. The disagreement of the unreasonable may bring them to play very counterproductive roles in society – the more unreasonable they are, the greater this risk. The state must take this into account and respond appropriately. Given the rationale for the state, responding appropriately means: acting in such a way that the negative impact of the unreasonable is minimized, and that perhaps some unreasonable people even ‘turn reasonable’.

If, then, theorists like Rawls and Valentini actually mean to say that the support of reasonable citizens is enough to justify the *implementation* of an institution (an interpretation that can, at least in Rawls’ case, not be ruled out, given the actual-world language that he uses), we ought to think again. For liberal rulers, when they are deliberating about institutions and policies, there are at least consequentialist arguments not to simply dismiss the opinions of those who are regarded as unreasonable from a (politically) liberal perspective.

The challenge for legitimacy theory is how to get from evaluative judgments about what is worth striving for to prescriptions about the implementation of institutions and policies. The fact that an institution or a policy is something to strive for is not normatively sufficient to justify its implementation. The reason for this echoes what I take to be, essentially, the same worry that Isaiah Berlin expresses when he says that “dogmatic certainty [about some final, perfect state] has been responsible for the deep, serene, unshakeable conviction in the minds of some of the most merciless tyrants and prosecutors in history that what they did was fully justified by its purpose” (Berlin 1958, 239). Ultimately, the adage “*salus populi suprema lex*” is a much better rule of thumb for rulers than “*fiat justitia, pereat mundus*” (cf. Galston 2010, 392).

In summary: protection against death and injury, security in the pursuit of our life projects, and increased opportunity for cooperation for mutual advantage are plausibly seen as values that it is the core purpose of the state to serve. When the state employs coercive power, it should be with an eye to realizing outcomes that serve these values. What this requires will differ according to the circumstances. Every context can put different demands on

the use of coercive power. This is the normative reason why a justification for coercive power should not be sought in evaluative tools, which are always to a greater or lesser degree context-independent. Rather, a justification should be given by showing how, in each situation again, state action contributes to the realization of the purpose of the state. We should, therefore, understand legitimacy in terms of circumstantial requirements rather than necessary requirements.

## References

- Berlin, Isaiah. 1958. "Two Concepts of Liberty." In *The Proper Study of Mankind: An Anthology of Essays*, edited by Henry Hardy and Roger Hausheer, 191-242. London: Random House. Original edition, 1997.
- Besch, Thomas M. 2013. "On Political Legitimacy, Reasonableness, and Perfectionism." *Public Reason* 5 (1):58-74.
- Buchanan, Allen. 2002. "Political Legitimacy and Democracy." *Ethics* 112 (4):689-719.
- Buchanan, Allen. 2004. *Justice, Legitimacy, and Self-Determination: Moral Foundations for International Law*. Edited by Will Kymlicka, David Miller and Alan Ryan, *Oxford Political Theory*. Oxford: Oxford University Press.
- Caney, Simon. 1998. "Liberal Legitimacy, Reasonable Disagreement and Justice." *Critical Review of International Social and Political Philosophy* 1 (3):19-36.
- Carey, Brian. 2015. "Towards a 'Non-Ideal' Non-Ideal Theory." *Journal of Applied Philosophy* 32 (2):147-162.
- Christman, John. 2005. "Autonomy, Self-Knowledge, and Liberal Legitimacy." In *Autonomy and the Challenges to Liberalism*, edited by John Christman and Joel Anderson. Cambridge: Cambridge University Press.
- Cohen, G.A. 2008. *Rescuing Justice and Equality*. Cambridge (MA): Harvard University Press.
- Erez, Lior. 2015. "Cosmopolitanism, motivation, and normative feasibility." *Ethics & Global Politics* 8 (1):43-55.
- Galston, William A. 2010. "Realism in political theory." *European Journal of Political Theory* 9 (4):385-411.
- Gaus, Gerald. 2011. *The Order of Public Reason*. Cambridge: Cambridge University Press.
- Geuss, Raymond. 2008. *Philosophy and Real Politics*. Princeton: Princeton University Press.
- Gibbard, Allan. 1976. "Natural Property Rights." *Noûs* 10 (1):77-86.
- Gilbert, Pablo. 2012. "Comparative Assessments of Justice, Political Feasibility, and Ideal Theory." *Ethical Theory and Moral Practice* 15 (1):39-56.
- Hampton, Jean. 1989. "Should Political Philosophy Be Done without Metaphysics?" *Ethics* 99 (4):791-814.
- Hobbes, Thomas. 1985. "Leviathan." In. London: Penguin Books.
- Larmore, Charles. 1996. *The Morals of Modernity*. Cambridge: Cambridge University Press.
- Larmore, Charles. 2013. "What Is Political Philosophy?" *Journal of Moral Philosophy* 10 (3):276-306.
- List, Christian, and Laura Valentini. 2016. "The Methodology of Political Theory." In *The Oxford Handbook of Philosophical Methodology*, edited by Herman Cappelen, Tamar Szabó Gendler and John Hawthorne, 525-553. Oxford: Oxford University Press.
- Locke, John. 1980. *Second Treatise of Government*. Indianapolis: Hackett Publishing Company. Original edition, 1690.

- Mack, Eric. 1990. "Self-Ownership and the Rights of Property." *The Monist* 73 (4):519-543.
- Nielsen, Kai. 2007. "There is No Dilemma of Dirty Hands." In *Politics and Morality*, edited by Igor Primoratz, 20-37. Palgrave Macmillan.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. Oxford: Blackwell Publishing. Original edition, 1974. Reprint, 2012.
- Otsuka, Michael. 1998. "Self-Ownership and Equality: A Lockean Reconciliation." *Philosophy & Public Affairs* 27 (1):65-92.
- Pettit, Philip. 2015. "Justice: Social and Political." In *Oxford Studies in Political Philosophy*, edited by David Sobel, Peter Vallentyne and Steven Wall, 9-35. Oxford: Oxford University Press.
- Räikkä, Juha. 1998. "The Feasibility Condition in Political Theory." *The Journal of Political Philosophy* 6 (1):27-40.
- Rawls, John. 1985. "Justice as Fairness: Political not Metaphysical." *Philosophy & Public Affairs* 14 (3):223-251.
- Rawls, John. 1993. *Political Liberalism*. Expanded edition ed. New York: Columbia University Press.
- Rawls, John. 1999a. *The Law of Peoples*. Cambridge (MA): Harvard University Press.
- Rawls, John. 1999b. *A Theory of Justice: Revised Edition*. Cambridge (MA): Harvard University Press.
- Reidy, David A. 2007. "Reciprocity and Reasonable Disagreement: From Liberal to Democratic Legitimacy." *Philosophical Studies* 132 (2):243-291.
- Sen, Amartya. 2004. "Elements of a Theory of Human Rights." *Philosophy & Public Affairs* 32 (4):315-356.
- Stilz, Anna. 2012. "Why Does the State Matter Morally?" In *Varieties of Sovereignty and Citizenship*, edited by Sigal R. Ben-Porath and Rogers M. Smith. Philadelphia: University of Pennsylvania Press.
- Timmerman, Peter. 2014. *Moral Contract Theory and Social Cognition: An Empirical Perspective, Theory and Decision Library A*: Springer International Publishing.
- Valentini, Laura. 2012a. "Assessing the global order: justice, legitimacy, or political justice?" *Critical Review of International Social and Political Philosophy* 15 (5):593-612.
- Valentini, Laura. 2012b. "Ideal vs. Non-ideal Theory: A Conceptual Map." *Philosophy Compass* 7 (9):654-664.
- Valentini, Laura. 2013. "Justice, Disagreement and Democracy." *British Journal of Political Science* 43 (1):177-199.
- Walzer, Michael. 1973. "Political Action: The Problem of Dirty Hands." *Philosophy & Public Affairs* 2 (2):160-180.
- Williams, Bernard. 2005. *In the beginning was the deed: realism and moralism in political argument*. Princeton, New Jersey: Princeton University Press.
- Young, Shaun P. 2006. "Rawlsian Reasonableness: A Problematic Assumption?" *Canadian Journal of Political Science* 39 (1):159-180.