



Report on the workshop “Ideal and Non-Ideal theories of Justice”: Towards a Non-Ideal Theory of Justice in Europe

Bert van den Brink, Simon Rippon, Tom Theuns & Miklos Zala

This Report was written within the framework of Work Package 2 “Philosophical Foundations for a European Theory of Justice and Fairness”

June 18



Funded by the Horizon 2020
Framework Programme of the European Union

Acknowledgements

First of all, we would like to thank our workshop participants, Ali Emre Benli, Oana Crusmac, Colin Hickey, David Jenkins, Hwa Young Kim, Aaron Landau, Sem de Maagt and Pedro Teixeira. As will be evident to those who know his work, the report is intellectually indebted in no small part to the work of Jonathan Wolff. We thank him for his keynote speeches in both the workshop and the plenary conference, and for his detailed comments in discussion. Frank Vandenbroucke, a member of the ETHOS advisory board, also gave very pertinent advice at the conference and workshop. The Center for Social Studies (CES), University of Coimbra were kind hosts of both the workshop and the first annual conference overall. We are also grateful to Eszter Kollar, Trudie Knijn, and Maria Paula Meneses for helpful comments on an earlier draft of this report.



Want to learn more about what we are working on?

Visit us at:

Website: <https://ethos-europe.eu>

Facebook: www.facebook.com/ethosjustice/

Blog: www.ethosjustice.wordpress.com

Twitter: www.twitter.com/ethosjustice

Hashtag: #ETHOSjustice

Youtube: www.youtube.com/ethosjustice

European Landscapes of Justice (web) app:
<http://myjustice.eu/>

This publication has been produced with the financial support of the Horizon 2020 Framework Programme of the European Union. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the European Commission.

Copyright © 2018, ETHOS consortium – All rights reserved ETHOS project

The ETHOS project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 727112

About ETHOS

ETHOS - Towards a European Theory Of Justice and fairness, is a European Commission Horizon 2020 research project that seeks to provide building blocks for the development of an empirically informed European theory of justice and fairness. The project seeks to do so by:

- a) refining and deepening the knowledge on the European foundations of justice - both historically based and contemporary envisaged;
- b) enhancing awareness of mechanisms that impede the realisation of justice ideals as they are lived in contemporary Europe;
- c) advancing the understanding of the process of drawing and re-drawing of the boundaries of justice (fault lines); and
- d) providing guidance to politicians, policy makers, advocacies and other stakeholders on how to design and implement policies to reserve inequalities and prevent injustice.

ETHOS does not merely understand justice as an abstract moral ideal, that is universal and worth striving for. Rather, it is understood as a re-enacted and re-constructed "lived" experience. The experience is embedded in firm legal, political, moral, social, economic and cultural institutions that are geared to giving members of society what is their due.

In the ETHOS project, justice is studied as an interdependent relationship between the ideal of justice, and its real manifestation – as set in the highly complex institutions of modern European societies. The relationship between the normative and practical, the formal and informal, is acknowledged and critically assessed through a multi-disciplinary approach.

To enhance the formulation of an empirically-based theory of justice and fairness, ETHOS will explore the normative (ideal) underpinnings of justice and its practical realisation in four heuristically defined domains of justice - social justice, economic justice, political justice, and civil and symbolic justice. These domains are revealed in several spheres:

- a) philosophical and political tradition,
- b) legal framework,
- c) daily (bureaucratic) practice,
- d) current public debates, and
- e) the accounts of the vulnerable populations in six European countries (the Netherlands, the UK, Hungary, Austria, Portugal and Turkey).

The question of drawing boundaries and redrawing the fault-lines of justice permeates the entire investigation.

Alongside Utrecht University in the Netherlands who coordinate the project, five further research institutions cooperate. They are based in Austria (European Training and Research Centre for Human Rights and Democracy), Hungary (Central European University), Portugal (Centre for Social Studies), Turkey (Boğaziçi University), and the UK (University of Bristol). The research project lasts from January 2017 to December 2019

Executive Summary

This report assesses the state of the art in ideal and non-ideal theory in political philosophy, and proceeds on this basis to make some practical methodological recommendations for integrating empirical and normative work toward a European theory of justice. It is based on the findings of the *Ideal and Non-Ideal Theories of Justice* workshop, held on February 23-24, 2018 at Coimbra University (Portugal), as part of the ETHOS annual conference. Abstracts of papers presented are included in Appendix 1. Selected working papers presented are included in Appendix 2.

A major part of the task of the ETHOS project is to bridge the gap between empirical and normative work. This raises the ‘is-ought problem’, identified by David Hume. Hume observed that normative propositions (e.g., ‘Domestic violence is wrong’) are entirely different from empirical ones (e.g., ‘Domestic violence occurs in every country’), and that it is thus unclear how the second could ever logically ground the first. This makes the idea of an ‘empirically grounded’ (normative) theory of justice appear paradoxical. However, the paradox is at least partially dissolved by noting that it is only empirical propositions *alone* that cannot generate normative conclusions. Accordingly, we advocate ‘hybrid empirical-normative theorizing’ in which empirical and normative work *collaborate* to generate the premises that ground normative conclusions about European justice.

The contrast between ideal and non-ideal theory concerns the ways in which theorizing about justice ought to be empirically informed. Hence, we have much to learn from philosophical debates on how theories of justice *can* and *should* be non-ideal. Following Valentini, we identify three main dimensions on which a theory can be ideal or non-ideal. The first is ‘full’ versus ‘partial’ compliance. (Ideal) full-compliance theory makes the idealizing assumption that persons in general accept and comply with the demands of justice. The second is (ideal) ‘fact-insensitivity’ versus ‘fact-sensitivity’ (e.g. sensitivity to empirical facts about human motivation). Finally, the third is the dimension between theorizing an (ideal) ‘end-state’ of complete justice, or articulating ‘transitional’ improvements towards greater justice. We note that transitional theory has an important advantage in light of *value pluralism*: while there is intractable disagreement about what full justice would be, we can often agree about gross *injustices*.

We thus advocate what we call, following Wolff, ‘real world political philosophy’, a bottom-up approach to justice in which one starts from manifest injustices. Voting rights for disabled persons can illustrate this approach well. We consider four models of disability: ‘medical’, ‘radical social’, ‘minority group’, and ‘human variation’. We show how these models suggest different remedies to the disadvantages faced by disabled persons, and highlight two concerns of justice that are salient: recognition and redistribution. In a case where the minority group model plausibly applies (e.g., a person with a mild mental impairment is denied voting rights) we argue that the solution is recognition based. In a case where the human variation model plausibly applies (e.g., a non-ambulatory person who cannot access public buildings to vote) we argue that redistributive considerations ought to be observed.

Real world political philosophy is ‘partial’; it does not provide a complete and systematic theory of justice or injustice, but rather focuses on identifiable injustices. This is a vitally important point for the ETHOS project, which responds to a grant call asking us to work towards a European theory of justice and fairness that rests on solid empirical ground. Our research and analysis so far indicates that *a* (unitary or monolithic) European theory of justice and fairness is not a feasible goal; the ETHOS project should rather direct its efforts toward developing *some* theory that helps us better understand, evaluate and recommend responses to European injustices.

Contents

Acknowledgments	<i>i</i>
Executive Summary	<i>iii</i>
<i>Towards a Non-Ideal Theory of Justice in Europe</i>	<i>1</i>
A. Introduction	1
B. Non-Ideal Theory and Theorizing (In)Justice in Europe	2
1. The Relevance of the Ideal/Non-Ideal Debates for Theorizing Justice In Europe	2
2. What is Non-Ideal Theory in Contemporary Political Philosophy?	3
3. Real World Political Philosophy as A Method of Non-Ideal Theory	6
4. Real World Political Philosophy: Applying the Model to ETHOS Research	10
C. Workshop Proceedings: Non-Ideal theory in Method, Conceptualization and Application	17
Bibliography	23
APPENDIX 1. Workshop Proceedings: Abstracts of Papers	26
APPENDIX 2: Selected Papers from the Workshop	36
Overview	36
1. Ideal Theory as Politics	37
2. The Importance of Ideal and Non-Ideal Theory	48
3. Feminist Political Theory and Non-Ideal Theory as its Methodology	66
4. Keeping Non-ideal Theory Real	80
5. Weaponising Justice and Equality	110
6. Ideal and Systematic Theory	133

Towards a Non-Ideal Theory of Justice in Europe

A. Introduction

As part of the first annual conference of the ETHOS project in Coimbra, Portugal, we organized an international workshop, *Ideal and Non-Ideal Theories of Justice*, on 23-24 February 2018. The keynote speaker, Professor Jonathan Wolff (Blavatnik School of Government, University of Oxford) spoke from his extensive experience at integrating normative-theoretical, empirical, and policy work, also in the conference plenary. The timing allowed workshop participants to attend main conference sessions, leading to a fruitful discussion on the challenges and possibilities of empirically-grounded normative theory.

The purpose of this report is threefold: first, after briefly exploring the relevance of the ideal/non-ideal debate for the ETHOS project, the current state of the art on ideal and non-ideal theory in contemporary political philosophy is analysed and discussed. This literature on non-ideal theory well shows both the challenges and the opportunities for an empirically informed theory of justice in Europe. The challenges from the philosophical perspective are, as we will see, centred on the classical methodological difficulty of moving from empirical observations of the world (such as are available to the social scientist) to the normative conclusions that a philosopher working in applied ethics seeks to draw. The problem is not limited to the domain of academic philosophy, however - it arises whenever a proscriptive or prescriptive proposal is made on the basis of empirical findings. This is, of course, central to the business of making or revising policy. Since policy recommendations are an integral part of the ETHOS project (especially, but not exclusively in WP7), tackling these philosophical-methodological issues will be invaluable for the project's success. Given that this report is intended to be of practical value to the other ETHOS work packages in how they pursue their goals, a balance has been made to review the most relevant recent work in the discipline in a manner that is rigorous, clear and philosophically productive.

The second goal of this report is to make practical methodological recommendations to the other work packages based on the research and investigations into the distinction between ideal and non-ideal theory. The aim is not, to be sure, to try to dictate to experts of disparate social scientific disciplines how to pursue research in their area of expertise. Rather the recommendations are pursued with a view to working towards the integration of the results of the various work packages in a manner that is coherent, harmonious and sound. Given the specialization of academic disciplines and their professionalization it is a great challenge to pursue interdisciplinary and multi-modal work. The structure of the ETHOS project includes philosophers at the preparatory stage as well as the stages of analysis and the integration of findings. In this report the idea is to bring epistemological and methodological research in political philosophy - explored during the workshop in Coimbra - in dialogue with the ways in which justice is conceptualized in the areas of law (WP3), political science (WP4), sociology (WP5), and critical political economy (WP6) respectively, as reported in the various x.1 deliverables.

The third and final aim of this report is to present some of the findings of the *Ideal and Non-Ideal Theories of Justice* workshop itself. This workshop included twelve presenters from ten institutions in seven countries. The keynote speaker was Professor Jonathan Wolff (Oxford) who spoke on 'real world egalitarianism'. Further

speakers included doctoral and post-doctoral researchers in political theory and political philosophy. The report will introduce the papers presented thematically, grouping them together around three poles: 1) papers centrally focused on developing methodological advances in the context of the ideal/non-ideal theory debate; 2) papers exploring and problematizing the distinction between ideal and non-ideal theory; and 3) papers developing ways to apply insights from the theoretical debate to ‘real world’ application. All the abstracts of the papers presented are included in the first Appendix. Finally, six of these working papers, by van den Brink, Crusmac, Jenkins, Kim, Landau, and Rippon and Zala have been included as a second Appendix. The content of the abstracts and working papers presented here is the sole responsibility of the individual authors.

B. Non-Ideal Theory and Theorizing (In)Justice in Europe

1. The Relevance of the Ideal/Non-Ideal Debates for Theorizing Justice In Europe

While not all of the debates in this area of philosophy have direct application to the work of the ETHOS project, some methodological aspects and controversies over the distinction between non-ideal and ideal theory do loom large. ETHOS sets as its task to develop and articulate an ‘integrative perspective on justice and fairness’ informed by the disciplines of political philosophy, sociology, law, economics, and political science. The way these disciplines function, however, can be strongly contrasting. As such, a core task is to ‘bridge the gap’ between empirical and theoretical work.

This bridging task engages many of the dimensions of the ideal/non-ideal debate in political philosophy. For instance, it is a theoretical question whether and to what degree normative theorizing ought to be ‘fact sensitive’, and which empirical facts, if any, are relevant to theorizing justice; however, once this has been determined, it is a matter of the empirical social sciences to determine the facts. To take an example, if we take seriously Simmons’ claim that certain incremental improvements towards justice (or away from injustice) may be morally impermissible due to their blocking circumstances of greater justice obtaining in the future (2010), we need empirical knowledge of what sorts of effects we ought to expect from various reform proposals.

A core philosophical puzzle that arises when pursuing this sort of interdisciplinary, bridging work between normative philosophy and empirical the empirical social sciences is the relation between the empirical ‘is’ and the normative ‘ought’. In his *A Treatise of Human Nature*, published in 1739, David Hume famously claimed that it the propositions ‘is’ and ‘is not’ are ‘entirely different’ from propositions ‘ought’ and ‘ought not’, a claim which is commonly referenced as the ‘is-ought problem’ (or sometimes as the gap between ‘fact’ and ‘value’). The ‘is-ought problem’ arises in the context of the ETHOS project since the ambition is to develop an empirically informed theory of justice and fairness for Europe. If empirical propositions and enquiry are analytically separate from normative propositions and enquiry, as Hume claimed, then the idea of an ‘empirically-informed’ theory of justice seems paradoxical, if not oxymoronic.

Some of the dilemma can be deflated however if we pay some more attention to the particular structure of the ‘is-ought’ problem; the problem, namely, is that we can’t move from ‘*plain*’ or ‘*brute*’ facts to normative judgements. Facts alone cannot themselves generate normative conclusions. In other words, an argument which has a normative (ought) conclusion, must have at least one normative (ought, or value-based) premise. Consider

the following example: person *A* promises freely to give person *B* some object *X* that they legitimately own. Many would say that it follows that person *A* *ought* to give *B* object *X*. This seems to be deriving a normative ‘ought’ (*A* ought to give *B* object *X*) from an empirical ‘is’, which the Humean objection seems to block. However, a better way to understand the is-ought problem is not to say that the conclusion is necessarily unfounded, but that the argument requires an unstated, *implicit* ought-type premise, along the lines that one ought ordinarily to keep one’s promises. Indeed, much work of applied ethics and normative political philosophy consists precisely of attempts to articulate the kind of implicit moral principles which guide or should guide our actions (or the regulatory principles of states) in particular circumstances.

It seems therefore we can say that an interdisciplinary project on an empirically informed theory of justice and fairness is possible, however not in the sense that the empirical work will generate all the ‘facts’ and the normative-theoretical work all the pre- and proscriptive normative conclusions and evaluations. Rather, empirical and normative work must *collaborate* to generate the premises generating valid normative conclusions about European justice and fairness. We can call this kind of structural collaboration ‘hybrid empirical-normative argumentation’. Insofar as this is the goal of (some of the) ETHOS work, the project is non-ideal in at least the sense that the normative conclusions that we aim at reaching are, clearly, ‘fact-sensitive’ (perhaps even *practice-dependent*). We return to this below, and see also Sangiovanni 2008, 2016).

2. What is Non-Ideal Theory in Contemporary Political Philosophy?

Already we have touched on the idea of ‘fact sensitivity’ as an important organizing idea for the debate over ideal and non-ideal theory. It will help to be a bit more systematic in exploring the ways in which non-ideal theories can be and have been taken to be non-ideal. First then, some taxonomy: the term non-ideal theory is used in contemporary political philosophy in several distinct though overlapping ways, almost always in opposition to ‘ideal theory’. One strand of literature focuses on trying to ascertain and reconstruct precisely what John Rawls - to whom we owe the original distinction - meant when he used the terms. It is in the light of this debate that the claim that something or other is *not* (non-)ideal theory is most often made (usually in the context of the author arguing that certain others misunderstand or misinterpret Rawls). Following usual practice and in the light of the ETHOS project’s broader goals, this assessment of the state-of-the-art focuses on ideal and non-ideal theories of *justice*, though much of what is written here will have application to ethical or normative theories more broadly.

A standard conceptual taxonomy of the ways theories of justice can be ideal or non-ideal distinguishes between three, sometimes overlapping metrics (Valentini 2012):

1. The degree to which a theory assumes full compliance
2. The degree to which a theory is ‘realistic’ or ‘fact-sensitive’
3. Whether a theory is ‘transitional’ or ‘end state’

Let us address these in reverse order of importance for the ETHOS project and start with the question of compliance. Roughly speaking, an ideal theory is sometimes taken to assume full compliance, while non-ideal theory is taken to assume imperfect compliance. This usage stems, as much of this debate, from Rawls’ *Theory of Justice*. In that seminal text, Rawls understands an ‘ideal theory’ of justice to operate under the presumption that, firstly, all persons who have duties of justice meet those duties fully and, secondly, that ‘natural’ conditions are

such that it is possible to realize justice, mainly in the sense that there are enough available resources for perfect justice to be possible (Rawls 1999a).

Clearly, even at this rough and general level, the degree to which a theory assumes full compliance will place it on a scale, so while there may be a ‘pure’ ideal theory (one that assumes perfect compliance) there won’t be a ‘purely non-ideal’ theory. If not purely ideal in the sense of assuming perfect compliance, a normative theory will be more or less non-ideal (incidentally, this is the case also for the question of the degree to which a theory is transitional and, especially, the degree to which it is ‘fact sensitive’). Stemplowska further separates the question of the degree to which normative theories assume less-than-full compliance into ‘permissive’ and ‘responsive’ parts, corresponding to different agents for which relevant non-ideal theories assume imperfect compliance (2017). ‘Permissive’ theories of imperfect compliance focus on the demandingness of claims on persons. Ideal theoretical models ought, on this view, not to assume that any individual should be fully compliant with their duties (of justice, for instance) because that would be too demanding *on them*. On the other hand, ‘responsive’ theories theorize imperfect compliance on the part of *others*; the idea here is that sometimes what one ought to do in the ‘real world’ is different than under idealized conditions of full compliance. Perhaps one ought to do more in the context of others’ less-than-perfect compliance, or less (Miller 2011), or perhaps one ought to do something wholly *different*.

In general, the Rawlsian ideal-theoretic move of (sometimes) assuming full compliance is less divisive and challenging than the two other senses that the distinction between ideal and non-ideal theory is often cast. Assuming full compliance is just one of many possible moves from the messy and complex reality to a *theory* which trades-off some degree of precision to achieve the sort of simplification (one could say ‘idealization’) required for speaking generally about normative demands across a wide variety of cases. Rawls was under no illusions that the assumption of full-compliance had some primary normative salience independently of its theory-building value. In other contexts, other simplifications will be more appropriate (*ibid*). Nor is it the case that pursuing ideal theory in the sense of assuming full compliance comes at the cost of, or is in any way a challenge to, pursuing non-ideal theory in this sense. It is also for this reason that it is not *necessary* either to assume full-compliance, or even to take a position on the usefulness of others assuming full compliance in their normative models when one pursues normative theory (that is non-ideal in this sense, and ideal or non-ideal in the other senses below).

The first conceptual dividing line between the ideal-theoretic assumption of full compliance and non-ideal theorizing in circumstances of less-than-full compliance overlaps to some degree with the question of the degree to which a theory is ‘fact sensitive’; after all, it is not very realistic to presuppose that others will fully comply with the demands of justice. However, the question of fact-sensitivity goes deeper than this. While Rawls generally operated ideal-theoretically in the *first sense* (assuming full compliance) in *A Theory of Justice*, he was famously criticized by G.A. Cohen for leaving the content of justice open too much to matters of human psychology, such as the propensity to avoid risk, or variable motivations to engage in productive activity (2008). Cohen thought, in other words, that Rawls’ theory was too ‘sensitive to the facts’ and ought to be *more* ‘utopian’ or idealized qua fact-insensitivity.

Consider Rawls’ *maximin principle* (or ‘difference principle’), which holds that justice demands that differences in material well-being are grounded on their maximising the position of the worst off. Rawls noted that his theory was agnostic between whether property-owning democracy or liberal socialism was to realize his principles of

justice in any given circumstance (2003). His view was that both were *prima facie* compatible with the demands of his theory, and that it was not the position of the political philosopher to ascertain empirically which specific set of politico-economic institutions would in fact maximize the position of a socio-historically defined group of ‘worst off’ persons. In sum, certain structural inequalities may be permissible in Rawls’ just society, provided that they are in the service of those who are worst off. In *Rescuing Justice and Equality*, Cohen contended, in contrast, that Rawls sometimes conflates justice with efficiency: while the latter may demand some trade-offs with perfect (‘ideal’) justice in devising practical rules of social regulation, they ought to have no impact on the articulation of principles of justice, which Cohen argued demanded a more thoroughgoing egalitarianism (2008).

Cohen went on to give a principled argument for why he thought that fundamental normative principles were always ‘insensitive’ to empirical facts in an article entitled ‘Facts and Principles’. The argument, crudely put, runs as follows: every normative principle X that seems to rely on a fact Y does so for a reason Z . Reason Z is itself a normative principle (Principle X_2), which either seems to itself rely on some empirical facts obtaining, or not. If Reason Z /Principle X_2 in turn seems to rely on fact Y_2 it must do so for a Reason Z_2 . To avoid an (impossible) infinite regress, all normative principles must eventually be grounded on a foundational Reason Z_n /Principle X_n , which is *insensitive* to facts (2003). Unsurprisingly, Cohen’s claim for the *necessarily* fact-insensitive (‘ideal’) nature of fundamental normative principles raised many objections; some prominent ones include that Cohen ignores that the choice of which political principles to endorse is subject to fact-sensitive political judgements (Hall 2013), that the logical validity of Cohen’s reasons mask the inapplicability of this reasoning to normative justifications (Jubb 2009, Ronzoni & Valentini 2008), or that Cohen neglects a crucial distinction between internal fact-sensitivity - whether a principle of justice references facts in its formulation - and external fact-sensitivity, understood as whether certain facts must obtain for a principle to apply (Pogge 2008). Setting aside these objections, it is important to note that even if fundamental moral principles are fact-insensitive, this does not preclude either reasoning about *non-fundamental* principles that *are* fact-sensitive, or non-ideal theorizing in either of the other two senses (assuming imperfect compliance, discussed above, or theorizing incremental or ‘transitional’ movements towards justice or away from injustice, discussed below). The meta-ethical problem does not, in other words, need to be solved for the ETHOS project to pursue an empirically informed theory of justice.

The third dimension often referred to by contemporary political philosophers and theorists through the language of (non-)ideal theory is the degree to which a normative theory articulates an ‘end-state’ of perfect justice, or focuses on articulating incremental, ‘transitional’ improvements towards more justice (or less injustice). Just as was the case with the first dimension of (non-)ideal theory discussed above (assuming partial vs. full compliance), nothing in the distinction between end-state and transitional theorizing about justice precludes the value of the other. It is perfectly plausible that both theorizing the end-state of perfect justice and theorizing what counts as incremental transitional steps toward justice are complementary tasks of political philosophy.

The literature on this matter is often framed through the opposition of these two ways of pursuing normative theory. This is largely a contingent matter of how the discipline of political philosophy developed in the late 20th century. Much Anglophone scholarship was dominated, once more, by John Rawls’ work on justice which, to a fairly large extent, operates ideal-theoretically with regard to its focus on articulating an ‘end state’. Mills has noted that Rawls’ ideal end-state theory is in a sense *doubly* ideal as Rawls both directs his theorizing to *an* end-state and idealized the circumstances leading up to this end-state in assuming no *historical* injustice (2017). Mills

labels this a distinction between ‘ideal ideal theory’ and ‘rectificatory ideal theory’, which would theorize an ideal end-state *to where we are now*, i.e. in conditions of manifest historical injustice (p.156-157). Like Mills, many scholars were unsatisfied with the putative dominance of end-state theory. Rawls’ work fuelled this response to a large extent; while he recognized the complementarity of end state and transitional theory, he claimed that transitional theory was in a sense parasitic on end-state theory. In *The Law of Peoples*, he wrote that ‘until the ideal is identified...non-ideal theory lacks an objective... by which its queries can be answered’ (1999b p.90; see also Simmons 2010).

The most prominent intervention on behalf of the value *and independence* of non-ideal theory was made by Amartya Sen. In an article ‘What Do We Want from a Theory of Justice’ (2006), and at greater length in his book *The Idea of Justice* (2009), Sen argued end-state theory (which he calls ‘transcendental theorizing’) is to a large extent redundant and even self-indulgent (cf. Robeyns 2012). He argued, *contra* Rawls, that one could identify justice improvements without a worked out ideal end-state by a method of comparison. The idea is intuitively plausible: it seems obvious that a world just like ours but without gender violence, for example, would be more just (or less unjust) than the world we are in. Furthermore, the seeming self-evidence of this position does not seem to rely on very substantive ideals about what the perfect end-state of justice would look like (beyond recognizing that it wouldn’t have any gender violence). Nevertheless, this common-sense position with respect to Sen’s view has been subject to wide criticism; Stemplowska for instance has argued that transitional theory is valuable but, *contra* Sen, benefits from end-state theory in that the latter gives a measure for determining the *extent* of ‘real world’ injustice (2017); Simmons has argued that the sort of ‘ordinal’ assessments of justice-improvements in Sen’s transitional theorizing requires knowledge of the end-state in order to be able to identify whether an incremental progression towards justice or away from injustice is morally permissible, especially in given it may sacrifice greater gains towards ideal justice in the long(er) term (2010, see also Stemplowska and Swift 2013). Rippon and Zala, in their working paper included here, argue that transitional theory can go badly wrong, mistakenly identifying comparative improvements, if ideal theory is neglected.

3. Real World Political Philosophy as A Method of Non-Ideal Theory

In the opening section, we noted how the model of having *only* empirical (or ‘brute’) facts as premises of an argument, but then ending with a normative conclusion, was philosophically suspect. We suggested that this could be addressed by pursuing *hybrid* forms of argumentation that includes both normative and empirical premises. To some extent however, this move to hybrid empirical-normative argumentation pushes the problem down the line. If we accept that a normative conclusion (either in the evaluative or the ‘action-guiding’ sense) requires at least one normative premise in the formal structure of the argument, then the question arises as to how to decide on and justify the appropriate normative premises that generate (in conjunction with empirical premises) the normative conclusions about justice and injustice, fairness and unfairness in Europe. This raises at least two issues. First, the manner in which the ETHOS project approaches the integration of political-philosophical theorizing about justice and fairness with empirical research seems to call for an ‘empirically informed’ (normative) theory of justice. The model above, assuming the ontological distinction between fact and value and the Humean is-ought problem, allows normative conclusions only with the inclusion of normative premises. If those normative premises *themselves* need to be ‘empirically informed’, then the same issue seems

to arise again as to the validity of this grounding relationship (between the empirical research - the facts - and the normative premises - the values). The second problem arises from the fact that there is not, as in some other disciplines, a general convergence in political philosophy around the accepted - or even the dominant - theory of justice, or fundamental moral theory. There is widespread disagreement about what theory of justice is correct both in civil society and in contemporary political philosophy. Furthermore, competing theories are often incommensurable. Isaiah Berlin calls the conjunction of disagreement about values and their incommensurability *value pluralism* (1969, see also Crowder 2002), while Jeremy Waldron talks about 'the *inevitable* existence of diversity of opinion and reasonable disagreement... on all matters of rights and justice' (1999 p.271, our emphasis).

One way to address the first problem would be to try to solve it, or to endorse a position which claims to. There are several noteworthy and serious philosophical positions which argue, in contrast to a sharp division between fact and value, that normative questions of political morality are always or often appropriately addressed in a manner sensitive to - even led by - empirical realities. To be alternatives to merely the integration of normative principles into a hybrid normative-empirical argument leading to a non-ideal normative conclusion, these responses stake a bold claim: the normative principles about politics themselves are reflexive to empirical circumstances. One prominent example of such an alternative is Andrea Sangiovanni's *practice-dependence* approach.

Sangiovanni's approach to theorizing justice and fairness in political contexts insists that the 'content, scope, and justification of a conception of justice depends on the structure and form of the practices that the conception is intended to govern' (2008 p.11 see also: 2016). If one wants to determine the appropriate conceptions of justice for a political institution, one must, therefore, determine what, functionally, the institution is for in the eyes of its practitioners. Interestingly, this approach seems to track in some senses functionalist approaches to solving the is-ought problem more generally; MacIntyre, for instance, claims in *After Virtue* that once one has determined the function of an object we can make evaluative claims about it (1981). For instance, using his example, it seems uncontroversial to claim that a watch is meant to keep time (that is its function). If a given watch is wildly inaccurate at time-keeping (fact-type observation) then MacIntyre wants to insist that it is not strange to conclude (value-type conclusion) that it is a *bad watch*. We don't need the additional normative premise that 'watches that fail to keep time well are bad watches' as that simply follows from identifying a given object as a watch. For all its interest and merits however, adopting Sangiovanni's practice-dependence approach or a functionalist analogue to determining ('empirically informed') normative inputs into hybrid empirical-theoretical arguments about justice in Europe would hold the conclusions hostage to the validity of this view, which, perhaps unsurprisingly, has raised much critical commentary (e.g. Valentini 2011, Meckled-Garcia 2013, Erman and Möller 2016).

This brings us, once more, to the second problem, disagreement - perhaps incommensurable, perhaps reasonable, certainly pervasive - about theories of justice and about fundamental moral theories. Regardless of whether the normative premise of a hybrid normative-empirical argument is itself 'empirically informed' (qua practice-dependant or otherwise) or not, it is extremely likely to be controversial. The more controversial the normative assumptions (or the theories generating them) that work as 'inputs' (premises) into arguments intended to generate, with empirical observations, normative conclusions about justice in Europe, the less stable and convincing those conclusions are going to be.

Perhaps an alternative method would be to map out several prominent normative theories, and to figure out what each of them would recommend in light of the relevant empirical findings? While this may be a worthy exercise in some cases (especially where, in a particular context, contrasting normative perspectives seem to converge on what they would recommend - see for instance Wolff and de-Shalit 2013), it doesn't always seem to be able to offer much by way of normative *evaluation*. Instead, it would generate a sort of normative 'menu', linking particular views (neo-republicanism, perhaps, or sufficiency-bounded luck egalitarianism) to particular policy evaluations or recommendations. The obvious question that arises then is *what to do* about injustice in Europe given conflicting normative conclusions. Further, there is no reason to suppose that the *right* normative theory is being included in the comparative approach (assuming there is one, ontologically speaking, despite reasonable epistemic disagreement over its content), especially if no work is being done to evaluate the theories that are called on. After all, the popularity of a normative theory is no measure of its validity since theories' popularities wax and wane over time.

Perhaps a more fruitful approach would be to focus the work not on highly controversial theories of ideal justice (and the consequent non-ideal evaluative and action-guiding conclusions generated when such ideal theories are put to work in hybrid empirical-normative arguments) but to already work at the input side with normative theories that are non-ideal in a certain sense. Recall the third sense that non-ideal was put to in the overview above: *transitional theory*. Theorists defending the possibility of transitional theory (Sen was discussed above) argue that it is possible to make incremental steps towards justice - or away from injustice - without having a picture in mind of perfect or ideal ('end-state') justice. The method that Sen develops is comparative: he thinks that people can come to a fairly good evaluation of justice by comparing two different states-of-affairs, on the one hand the current state-of-affairs in a given circumstance, and on the other the expected state-of-affairs after a reform or policy change (cf. Simmons 2010). Sen's approach is most powerful and convincing when it is used modestly, i.e. when the real-world state-of-affairs in question is grossly or *manifestly* unjust in a way that the reform or policy change is expected to alleviate. The earlier discussion of Sen's idea used the example of gender violence: it just seems obvious, even in the absence of an ideal (end-state) theory of justice, that decreasing the instances of gender violence would reduce the level of injustice in the world, other things being equal (cf. Simmons 2010).

One way to understand this modest version of Sen's comparative approach is to re-evaluate the idea of value pluralism in light of gross injustice. The idea is, basically, that accepting that there is reasonable disagreement and, possibly, value pluralism about *perfect justice*, does not mean we need to be committed to such reasonable disagreement about *gross injustice*. Perhaps reasonable views can and do converge when confronted with such injustices. If so, it would not even matter, methodologically, what, philosophically, grounds the injustice in question (presumably this *would* be the object of widespread disagreement) if one would want to articulate practical proposals for alleviating injustices. Consider for example the idea that it is unjust for states not to reasonably accommodate disabled persons to enable them to exercise their right to vote. That this is unjust may be (reasonably) grounded on a plurality of different reasons: it may be considered unjust qua a violation of disabled persons' human rights, it may be considered unjust qua in conflict with procedural democratic justice, it may be considered unjust qua unfairly discriminating between disabled and able-bodied people, and so on. Some of these reasons may be congruent, while other reasons may be incommensurable. Still, this view holds: regardless of the disagreement over *why* it is unjust, we can agree that it *is* unjust for states not to reasonably accommodate disabled persons to enable them to exercise their right to vote. A state that reforms the

organisation of elections and electoral policy would then, in this precise sense, become less unjust. In other words, more generally stated, we don't need to settle the precise nature of *why* certain injustices are unjust to be able to identify them as plausibly unjust.

This approach to prioritizing gross or manifest injustice may seem to undermine the interest in the kind of hybrid empirical-normative theorizing that is being supposed. A critic may wonder why, if injustice is *manifest*, much is needed by way of empirical or indeed normative research. This position though is wrongheaded. It is not because certain injustices are manifest *when they are in view* that they are obvious in the absence of careful analysis. Many manifest injustices may be hidden from sight, either because of a lack of attention to a particular issue, or because there are reasons for them to be hidden (both in the passive and the active sense). In a 1990 article, Sen shows that there are 100 million fewer women alive than ought to be expected by biology, a difference that cannot be (fully) explained by selective abortions. This leads to the conclusion that millions of girls and women are dying earlier than they should, relative to boys and men, due to neglect (1990, see also Wolff 2015). The idea that this situation is a *manifest injustice* does not discount the fact that it took painstaking and difficult work to recognize that it was in fact the case. This can be generalized: it will often take painstaking and careful empirical work to determine 'the facts', some of which, when in view, will appear manifestly unjust. We could call this sort of work, a little clumsily, *non-ideal hybrid empirical-normative theorization of injustice*. More simply we suggest adopting Jonathan Wolff's phrase 'real world political philosophy' to describe this approach (see below).

A political philosopher that defends a related view is David Wiens. In an important article in the *Journal of Political Philosophy*, Wiens argues in favour of a type of non-ideal theorizing that he calls 'clinical theorizing' (2012). This is characterized by what he calls an '*institutional failure analysis approach*' that sees its main task as 'obviating or averting social failures' (p.46). What this approach entails, is starting with identifying a failure of an institutional set of arrangements, for instance a gross injustice or an instance of unfairness, then diagnosing the cause of the failure, and finally designing alternative institutions that avoid that can be expected to avoid that particular failure. Now it seems that this approach still relies on 'ideal theory' and 'values' in a sense, since to identify something as a failure one needs to be able to refer to a metric of value on which it fails. However, Wiens holds that there is a distinction between *ideal theory* in the senses explored above, and a 'theory of ideals', or values (see also, centrally, Hamlin and Stemplowska 2012). Arguably, the type of non-ideal theorizing Wiens proposes (similarly to the comparative approach that Sen develops) *does* need to make some reference to ideals and values (namely ideals and values about what in a particular social and political institutional context constitutes a failure), but it does not need to assume full-compliance and cannot be, in any straightforward sense, fact-insensitive or targeted to an end-state, and therefore does not presuppose ideal theory (cf. Simmons 2010).

Jonathan Wolff (the keynote speaker of the WP2 workshop on Ideal and Non-Ideal Theories of Justice) makes a similar move to prioritizing a certain kind of 'failing', focused on social inequality (2015). Like literature on ideal justice, the question of what ideal equality would look like is highly disputed in contemporary literature. Wolff's idea however is that there is in a sense an asymmetry between the disputatious nature of conceptualizing ideal equality, and the convergence of views of what constitute inequality. Social egalitarians can therefore 'have a clear sense of what they are against - hierarchy, snobbery, servility, oppression' without spelling out a perfect vision of what they are for. What they are for, more simply understood, is merely those social relations, in the plural, that are characterised by the *absence* of social inequality. Wolff argues that there is no *logical* reason to presume that the positive notion ('equality', but equally 'justice') must have conceptual priority over the negative

notion ('inequality', 'injustice') in the sense the second must derive from the first - this is merely an accident of grammar (see also Shklar 1990). He further holds the view, on which we need take no position here, that this means that the subsequent *range* of models of social organization (namely those that avoid social inequality) cannot be argued to be better or worse from one another on social egalitarian grounds and that any philosophical argument for preferring one over the other would need to use other grounds to distinguish between them.

4. Real World Political Philosophy: Applying the Model to ETHOS Research

Whereas idealizations are essential for political philosophy (even for non-ideal theory, see Rippon and Zala's paper in Appendix 2), we have argued that if we want to utilize philosophy for the purpose of public policy, we had often better apply a non-ideal theoretical, bottom-up approach to justice that starts from real world problems or injustices and aims to provide solutions to them (see Wolff 2011). Contrary to an ideal theoretical approach that would aim to identify and clarify values first and/or construct abstract principles, the bottom-up view starts from a diagnosis. That is, it examines how things go wrong in the real world instead of what would justice require in an ideal situation. Real world political philosophers work in that sense in a manner akin to a clinician who examines the patient first and then tries to provide a cure to the patient's problem.

An important feature of this mode of theorizing is that it is partial as it does not aim to provide a complete account of a just society. It either focuses on a given real world problem, such as immigration (Cole and Heath Wellman 2011), the regulation of drugs (Husak and de Marneffe 2005), gambling (Wolff 2011 Ch. 2), prostitution (de Marneffe 2010), or same-sex marriage (Corvino and Gallagher 2012), just to name a few, or it analyses the social mechanisms that bring about injustice. The approach we propose here for how to *do* non-ideal normative theorizing in an empirically-informed manner, at the 'input stage' of hybrid empirical-normative argumentation about *injustice* in Europe, is congruent to the approach favoured by Nancy Fraser, whose schema of justice as redistribution, recognition and participation (representation) has been adopted by the ETHOS project. The following passage from Nancy Fraser illuminates this feature of non-ideal theorizing:

For me, a theory of justice cannot, and should not, provide a comprehensive account of the overall goodness or badness of society. Rather, it should allow us to evaluate social arrangements from the perspective of one limited, but extremely important angle: how fair or unfair are the terms of interaction that are institutionalized in the society? Does the society's structural-institutional framework, which sets the ground rules for social interaction, permit all to participate as peers in social interaction? Or does it institutionalize patterns of advantage and disadvantage that systematically prevent some people from participating on terms of parity? Do the society's institutionalized patterns of cultural value create status hierarchies, which impede parity of participation? Does its economic structure create class stratification, which also forecloses the possibility of parity? (Fraser et al. 2004, 367)

We can find similar analyses in the work of Elizabeth Anderson (2010) and Iris Marion Young (2009). Part of what is needed in the current project of justice in Europe is such a bottom-up, problem solving approach to justice. This requires a division of labour between empirical and normative research. For bottom-up theorizing requires plenty of empirical information to be able to get off the ground (cf. Wolff 2011, Ch. 9).

It is important to clarify that when we advocate a bottom-up, partial view of justice, we do not thereby propose a view that is ‘isolationist’ in Simon Caney’s sense (2012, 258-9). By ‘isolationist’, Caney means a treatment of the justice of one issue (e.g. climate change or gender) that treats it as if it were independent of all other issues of justice, and of general principles of justice. In the bottom-up approach we advocate, as Jonathan Wolff writes, ‘the first task is to try to understand enough about the policy area to be able to comprehend why it generates moral difficulties, and then to connect those difficulties or dilemmas with patterns of philosophical reasoning and reflection’ (Wolff 2011, 9). This approach need not be isolationist, since one can go on to connect bottom-up philosophical reflection on one real world problem with broader normative considerations, including those of a more comprehensive political theory. In the end, while we believe that this integrative philosophical work is valuable, it does not follow that this further work must be provided by the bottom-up, partial theory itself.

Because of its attention to specific, identifiable injustices, not only is non-ideal theory of the kind we have been advocating partial (in the sense that cannot hope to provide a complete normative theory of injustice, analogous to an ideal account of the just society), it is also rather difficult to characterize in general terms. Perhaps there is no better way to characterize this kind of theorizing than by example. Therefore, we now turn to an example of the kind of non-ideal, real world philosophy of injustice we have in mind, in order to demonstrate this approach. This could be used as a model for interrogation of the other issues of injustice that ETHOS investigates. Here, we focus on aspects of ETHOS’s Work Package 3, which examines European law on *voting*, *housing* and *education* in the particular context of vulnerable groups such as *refugees*, *people living with disabilities*, and *ethnic and religious minorities*. We will focus on just one of these groups: the disabled. We will indicate how in the real world justice concerns of recognition and redistribution arise in connection with this group, and indicate how we might better conceptualize these concerns of justice so that they can be clearly identified when thinking about legal regulation.

a) Models of Disability

What makes disability interesting from a political philosophical point of view is that it involves two fundamental aspects of justice: recognition and redistribution (see Wasserman et al. 2013, Section 1). As far as the former is concerned people living with disabilities are frequently victims of exclusion and stigmatization. As far as the latter is concerned, disability is a problem because disabled individuals are often disadvantaged along certain socioeconomic dimensions (such as income, wealth, or opportunities) (Wasserman et al. 2013). This twofold concern evokes Nancy Fraser’s (1995) two-dimensional framework of justice as *recognition* and *redistribution*. Indeed, as Gideon Calder points out, ‘in particular, internally diverse ways, people with disabilities have been on the end of a kind of pincer movement between Fraser’s two impediments to parity; maldistribution and misrecognition’ (2010, 62, quoted in Wasserman et al. 2013).

Thus, justice for the disabled can be conceived as a requirement that people with physical/mental impairments should not suffer disadvantages due to their atypical physical or mental characteristics. This needs an approach that we described above: one that *first* aims to understand the ways in which people with physical/mental impairments will be disadvantaged via the social and physical environment and which *then* provides a solution how to counterbalance those disadvantages. In other words, we need a diagnosis of the injustices suffered by the

disabled first and *then* we can think of how to provide remedies to address these injustices. In other words, in the method of Real World Political Philosophy, we start by examining the different explanatory models of disability that aim to explain and interpret the phenomenon of disability. The next step is to use these observations to help determine what types of remedies are called for.

Mainstream models of disability are of four different types: A) the medical model, B) the radical social model, C) the minority group model, and D) the human variation model (Wasserman et al. 2016). These models provide different explanations how disability works and why it is a problem.

(1) The Medical Model

The medical model of disability is the way many health care professionals tend to look at the issue: disability is an individual pathological condition that results in certain kinds of personal and social limitations. This model ‘sees the limitations faced by people with disabilities as resulting primarily or solely from their impairments’ (Wasserman et al. 2016). The WHO’s 1980 International Classification of Impairment, Disability, and Handicap seems to conceive disability in this way. While the document acknowledges the social aspect of disability, it defines disability (in the text referred as handicaps) as ‘disadvantage experienced by the individual as a result of impairments and disabilities’ (WHO 1980, 14; Altman 2001, 109).

Following Carol Thomas, the medical model of disability can be described as understanding disability as either caused by impairment, or as one and the same thing, i.e. (Impairment → Disability), or (Impairment = Disability) (Thomas 1999, 14). In other words, the medical model holds that whatever disadvantages physically/mentally impaired persons face at the social level necessarily stem from their impairments.

(2) The Radical Social Model

The medical model of disability has become the target of disabled rights movements since the 1960s. They embraced the *social model* of disabilities, according to which disabilities are the results of the exclusion and discrimination of the mainstream society. The Union of the Physically Impaired Against Segregation (UPIAS), which was the leading organization of the social model in the UK went so far to deny any causal role to physical impairments in creating disabilities. In a seminal text, the activists of UPIAS stated: ‘In our view, it is society which disables physically impaired people. Disability is something imposed on top of our impairments, by the way we are unnecessarily isolated and excluded from full participation in society’ (UPIAS 1976, 3, quoted in Shakespeare 2006, 29). Thus, the social model holds that it is the mainstream society that creates disability from personal physical/mental impairments; in Thomas’s diagram: (Social barriers → Disability) (Thomas 1999, 14).

(3) The Minority Group Model of Disability

Today few people support the pure versions of these first two models; most theorists position themselves somewhere between the two. As Wasserman et al. point out, the medical model is ‘less often explicitly defended than unreflectively adopted -- by health care professionals, bioethicists and philosophers who ignore or underestimate the contribution of social and other environmental factors to the limitations faced by people with disabilities’ (Wasserman et al. 2013). They also point out that there is a ‘growing awareness of the environmental contribution to disability and a partial embrace of the social model’ (Wasserman et al. 2013).

Thus social model theories became more nuanced as they emphasized the *interaction* between individual characteristics (impairments) and the social/material environment (Wasserman et al. 2005). Two contemporary social models merit closer consideration. The first is the minority group model: emphasizing the discrimination of mainstream society, the minority group model of disability holds that disabled people are relevantly similar to other disadvantaged groups, such as racial or ethnic minorities, and physical or mental impairments are among the suspect categories of discrimination law. The minority group model is very popular among disability scholars (for some of its notable proponents, see Silvers 1998; Gliedman and Roth 1980).

(4) The Human Variation Model

The other version of the social model, the human variation model, holds that the main problems of the disabled do not stem from the deliberate or indirect discrimination of society, instead, the problem is the *mismatch* between the characteristics of people with physical/mental impairments and the social/physical environment (Wasserman et al. 2013).

According to its main proponents, Richard Scotch and Kay Schriener:

[d]isability could be defined as an extension of the variability in physical and mental attributes beyond the present – but not the potential – ability of social institutions to routinely respond... The problems faced by people with disability might be seen as the consequence of the failure of social institutions (and their physical and cultural manifestations) that can be attributed to the institutions' having been constructed to deal with a narrower range of variation than is in fact present in any given population (Scotch and Schriener 1997, 155).

In other words, the human variation model does not consider the disabled as an identifiable minority group, but instead views people living with disabilities as 'individuals who simply differ as a matter of degree from population norms for one or more physical characteristics' (Wasserman et al. 2013). Hence, looking at disability through the lens of human variation, the problem is that one size does not fit all and that the physical/social environment cannot take every potentially occurring individual variation into account.

b) What to Do with Disability: Remedies

Because they give different diagnoses of what disability consists in, these different models also suggest different remedies. Anita Silvers (1998, 16) enumerates three possible responses to the problem of disability:

1. We provide medical treatment and rehabilitation that restores people's functionality.
2. We manufacture adaptive equipment such as talking computers and power wheelchairs that assist people's functionality.
3. We alter the environment to make it more functional for people who cannot see, hear, walk, or execute other familiar performances.

In a similar fashion, Jonathan Wolff (2002; 2009) proposes the three following possible solutions: 1) *Personal enhancement* (such as surgical interventions) that develops someone's functional capacity, 2) *targeted resource enhancement* (such as giving someone a wheelchair to be able to get around) that provides an external resource

to the person to restore a functional capacity, and 3) *status enhancement*, that requires changing the environment so as to help someone regain her functionality (e.g., designing buildings with ramps to make them accessible for non-ambulatory wheelchair users). There is a fourth possible response: *compensation* (either monetary or in-kind) (cf. Wolff 2002). However, in most cases this is not a morally acceptable solution – as Jonathan Wolff points out, ‘I do not know of an argument from within the disability movement that the special miseries of the disabled need cash compensation, and no doubt this would be considered deeply insulting’ (Wolff 2002, 211). Or, as Ronald Dworkin puts it, if corrective technology is not available, merely swelling the bank account of a person who was born blind is not a proper solution (Dworkin 1981, 299).

Having these possible responses in mind, we can see the different diagnoses and cures of the three respective models (see Wasserman et al. 2013). The medical model sees the problem in personal (physical or mental) impairment, and suggests the cure of personal enhancement. The medical model can also recommend targeted resource enhancement and status enhancement, but only as second-best solutions. The radical social model sees the problem in disabling environments and suggests their reconstruction (this model disfavours personal enhancement as a solution). The minority group model holds that the problem is the exclusion and discrimination of mainstream society, which suggests as a remedy better recognition (status enhancement) for the discriminated against group. The human variation model also holds that the problem is with exclusive environments and suggests the remedy of reconstruction of the environment, but the emphasis here is not on discrimination, but on the problem that the given social or physical environment has a limited capacity for accommodating atypical personal characteristics.

c) Concerns of Recognition and Redistribution

At this point we can return our twofold justice concerns of recognition and redistribution. Recognition will be an integral part of the above-mentioned social models. Recognition of the disabled will require an end to social exclusion and stigmatization of the disabled. Thus, giving recognition to people with disabilities often requires adopting the perspective of the minority group model. But recognition will also demand the perspective of the human variation model: we should provide accommodations to the physically and mentally impaired by tailoring the environment to their characteristics left out by the current design of mainstream social arrangements. Whereas recognition will entail resource and monetary reallocations, this is not a concern of redistribution so far. As David Wasserman points out,

Consider... the question of how much it is reasonable for a small business to spend on an elevator or ground-floor space to be able to employ a talented IT technician with emphysema. To answer that question, we might do better to decide what respect for that person demands, based in part on current social practice and convention, rather than to consult a comprehensive but abstract theory of distributive justice (Wasserman 2014, 271).

But concerns of *redistribution* will also be essential, because providing personal enhancement, targeted resource enhancement, or status enhancement is often significantly costly and this cannot be indifferent to any plausible theory of justice (nor to policymakers, for that matter). Consider the legal notion of reasonable accommodation, which is a crucial part of both the 1964 Civil Rights Act (concerning religion) and the Americans with Disabilities

Act (concerning disabilities). These legal acts declare that it amounts to discrimination if religious and disabled individuals are not 'reasonably' accommodated in their workplace, i.e. if they are denied accommodation that would not present an 'undue hardship' to the employer.

To use this legal notion, redistribution in the case of the accommodation of disability aims to clarify the limits of accommodation. For example, if costs matter from the point of view of justice then some disadvantages of the disabled will be justified, such as in the case of not making a seventeenth century building accessible for people using wheelchairs at enormous cost (cf. Wasserman et al. 2013). But those disadvantages that the non-accommodated have to endure are manifestly unjust if accommodation would not cause an unreasonable hardship to those under a duty to accommodate.

What is unavoidable is a moral cost-benefit analysis, wherein both the majority and the minority need to provide reasons why they want to uphold or change a given social arrangement. From the point of view of the majority, the efficiency of the current arrangement and the costs of change are certainly factors, whereas the minority which desires accommodation has to be able to show that their burdens are unjust and not mere inconveniences.

This balancing between costs and inclusion is aptly exemplified by a metaphor of Allen Buchanan (1996). He illustrates the problem of disabilities by a card game: the more complex the rules of the game that we choose to play, the fewer people will be able to participate in it. That is, if we choose bridge then some individuals (e.g., toddlers) will not be able to play (Buchanan 1996, 40). If we choose a less complex game, more people will be involved, but if we choose an extremely easy one, such as Go Fish, then many will not be interested in playing the game. According to Buchanan, there are two legitimate interests at stake, one is inclusion, the other is having the most effective social arrangement available.

More generally, it is important from a normative point of view if a given social arrangement can be more inclusive, but the efficiency of that arrangement is also important for justice, and the two considerations must be balanced. To give a hypothetical example, whereas there would be definitely some value in cushioning the sides of roads so as to make driving available to people with severe vision impairments, it could be very ineffective if traffic would become significantly slower (and perhaps even unreasonably costly). Thus, justice demands the reconstruction of social arrangements with the goal of inclusivity only to the extent that these can be achieved at a reasonable cost, and only insofar as they would not make the given social arrangement unreasonably ineffective. Such trade-offs, of course, are always context dependent, but they are unavoidable.

To conclude: In order to be able to justify the accommodation of the disabled, legal frameworks should apply a two-stage analysis. First, they must start with a diagnosis. Second, they must clarify what kind of response justice requires in the given situation. To provide a diagnosis, one must start from the inquiry of what kind of personal impairment or characteristic must be accommodated and why the given impairment or characteristic is a problem in the first place. Disability models can do this explanatory work. In the second stage, if there is a proper explanation of the given case in hand, one has to find out which response to disability fits best in the given case, and what justice considerations motivate the choice. So, should we change the person's characteristics? Should we provide targeted resource enhancement, or change the social or physical environment instead? Or should we provide some kind of compensation, for example if the first three responses are not available? This choice will be informed and motivated by the normative concerns of recognition and redistribution (specifically, avoiding manifest *misrecognition* and *maldistribution*). It is worth starting with the concern of recognition, that is,

the question what would a given condition require to be treated with respect, and then clarify whether there are limits to accommodation justified by the concern of redistribution.

To illustrate this two-stage method with a more specific example, consider two individuals with two different kinds of impairment: one with a mild mental impairment and one a non-ambulatory person. Suppose that, as policymakers, we want to make our electoral system more inclusive and we want to find out whether voting rights for these two individuals (and people like them) should be accommodated by the new system, or not.

First, starting with the diagnosis, we must examine what kind of challenges someone faces if they want to cast a vote, and whether those challenges are necessary and justified parts of election procedures, or not. So, voting requires certain mental capacities but it should not require physical capacities – being able to vote should not depend on someone’s physical fitness. Thus, it is arguable that the exclusion of someone with a severe mental impairment from voting is justified, but it would be clearly unacceptable to deny a wheelchair user’s right to vote on the grounds that they cannot walk. Similarly, if the mental impairment is not sufficiently severe, it might be unacceptable to deny the right to vote of a person with such a limitation.

Continuing with the diagnosis: we have find out which model of disability provides the most plausible explanation of the case. What is the nature of the problem? Imagine that there is person with a mild mental impairment but who has more than the minimal rational capacity to be able to meaningfully choose among the available political options. But they cannot vote, because it is forbidden by a law motivated by false stereotypes about the mentally disabled. The minority group model seems a good fit for this case as an explanation -- the medical model clearly cannot help here because whereas personal enhancement of the mentally disabled voter could help them to ‘fit in’ this would not be an adequate response, morally speaking. To see this, consider an analogous situation regarding racial discrimination. The medical model approach would seem to suggest that an appropriate ‘solution’ to racism would be to change a black person’s skin colour so that they may be able to enter to a public facility that discriminates against blacks. Clearly, such an approach is grossly unjust and offensive. So, to bring the argument back to disability, the person with a mild mental impairment might ask for the elimination of legal hurdles. Imagine also that the non-ambulatory person faces the challenge that voting booths are in non-accessible public buildings. There are no legal hurdles (wheelchair users are not excluded from the franchise), but there are clearly social barriers. These barriers are properly explained by the human variation model: the problem is that public buildings take only a limited range of personal characteristics into account (they are not designed to accommodate the needs of non-ambulatory people).

The next step is to consider in normative terms the possible responses to these diagnoses. Status enhancement seems to be a proper first response in both cases, because recognition requires accommodation in both cases. But considerations of redistribution can qualify our response. For example, whereas eliminating laws that discriminate against people with mild mental disability is costless, reconstructing public buildings might be very expensive. It is possible then, that providing tools for remote voting in the form of mobile voting booths is a more cost-effective solution, thus a targeted resource enhancement would be a cheaper option that would still satisfy recognition concerns. Policymakers must think through similar considerations regarding efficiency (would a given accommodation slow down the elections too much? And so forth).

These hypothetical examples are of course by no means exhaustive, they simply aim to illustrate the kind of non-ideal theoretical analysis that normative conclusions (and, by extension, policy recommendations) concerning

disability accommodation requires in the context of voting. The same analysis is unavoidable in the context of housing and education, and also in case of other groups such as religious and ethnic minorities or refugees which are in the focus of WP 3.

We earlier described real world political philosophy, or non-ideal theory, as ‘partial’, in the sense that we doubt it can provide a complete and systematic theory of justice or injustice, but rather ought to focus on the analysis of identifiable injustices. We hope that the preceding short illustration of how a philosophical analysis of disabled voting rights might be developed in real world political philosophy makes this point clear. This is a vitally important point for the ETHOS project, which responds to a grant call asking us to work towards a European theory of justice and fairness that rests on solid empirical ground. Our research and analysis so far indicates that *a* (unitary or monolithic) European theory of justice and fairness is not a feasible goal; the ETHOS project should rather direct its efforts toward developing *some* theory that helps us better understand, evaluate and recommend responses to European injustices.

C. Workshop Proceedings: Non-Ideal theory in Method, Conceptualization and Application

It is unsurprising that many of the papers presented at the workshop have a methodological slant as the debate between ideal and non-ideal theory in political philosophy has, as we have seen, centred on issues of method. Four of the contributions, by de Maagt, Rippon and Zala, van den Brink and Wolff, could be characterized as having a predominantly methodological focus in the sense of theorizing how political philosophy can and ought to approach ‘applied’ normative analysis. We have seen that some, radical, positions in the state-of-the-art in part one of this report saw little space for the systematic normative theorization of applied political problems (Miller 2013, chap 11). We can call this position the ‘sceptical position’; clearly, if the sceptical position is correct, there is very little room for an empirically informed theory of justice such as the ETHOS project seeks to develop. It is thus heartening that there is little support for the sceptical view amongst the researchers at the workshop in Coimbra.

The five authors grouped here under the banner ‘methodology’ all disagree with the sceptical view, albeit in different ways and to differing degrees. De Maagt seeks to develop an alternative account to the dominant approach of applied theorizing in contemporary political philosophy, which focuses on reaching a ‘reflective equilibrium’ or coherence, between principles and normative judgements. In its place de Maagt sketches out a ‘transcendental approach’, inspired by the philosophy of Immanuel Kant, which tries to show that certain sorts of moral claims cannot be rejected without self-contradiction by agents engaging in communicative interaction with others. De Maagt’s transcendental approach starts with articulating a supreme principle of ethics - that everyone has the right to the ‘necessary means of agency’ - and applies this principle in a hierarchical fashion to practical (non-ideal) questions. In the context of ETHOS, such an approach would seem to require that the first step of the project is a full and thorough articulation of a (transcendental) moral theory of justice *methodologically prior* to obtaining empirical results which, in turn, would be prior to analysis and policy recommendation. In this sense, de Maagt calls for an approach to non-ideal theory which takes ideal theory as its starting point and reference.

Simon Rippon and Miklós Zala argue that, since we want political theory to help solve the very practical problem of what to do, non-ideal theory is unavoidable (this contribution is in Appendix 2). Nevertheless, they insist on the importance of ideal theory for the business of non-ideal theorizing. They take as a target Sen’s ordinal approach

to justice, which emphasizes the importance of comparative judgments on the relative justice or injustice of certain actual states-of-affairs. While endorsing, to a certain extent, Sen's view that the comparative approach allows one to make comparative assessments of justice and injustice without ideal theory, Rippon and Zala also seek to defend the plausibility and value of an ideal-theoretical approach. They argue, in effect, that any comparative justice judgements must make use of at least a 'vague ideal of the just', which ideal theory can help to refine. While de Maagt's approach to non-ideal theory can thus be seen as subordinate to ideal theory, Rippon and Zala conceive of the relation as a symbiosis. Given ETHOS' inductive approach to theorizing justice and fairness, their paper does serve as a warning about the reach of the methodology for generating normative judgements about justice in the absence of ideal theorizing.

In the context of political pluralism, Bert van den Brink critically questions the plausibility of achieving a singular ideal theory of justice (this contribution is in Appendix 2). He took his inspiration methodologically from Quentin Skinner's important contribution to the debate on methodology in the history of political thought (Skinner 1969), and from the group of political theorists and historians of political thought collectively known as the 'Cambridge School'. They argue, and demonstrate, that many of the thinkers that are typically thought to have furnished 'foundational' theories that address 'timeless' questions of politics in fact are better understood as intervening in actual, historically bounded debates on practical questions of politics contemporary to their time. Van den Brink argues that a similar process of reasoning ought to make philosophers question the timelessness and foundational character of their own philosophizing about justice and fairness, which itself springs, of necessity, from the political context of the day. Nevertheless, he argues that this did not mean that political philosophy could do without ideal theory, defending the utility of philosophers working out separate (and sometimes disparate) ideal theories of justice. Conversely, such self-awareness of the boundedness of ideal theory can serve both to improve the systematization and coherence of ideal theory on the one hand and, on the other, furnish the practical philosopher with a more diverse toolbox of theories that can be mobilized in the face of problem-oriented non-ideal theorizing. As such, van den Brink's account lends credence to a somewhat theoretically eclectic and problem-driven approach to non-ideal theorizing, something we shall see come back also in Jonathan Wolff's argumentation.

Jonathan Wolff opened his keynote lecture by focusing on the choice of terms 'ideal' and 'non-ideal' theory. He argues that the latter term ought to be rejected, as it makes it appear parasitic on ideal theory. Wolff proposes that the term 'real world political philosophy' could be used to replace 'non-ideal theory'. The notion of a society of equals is, Wolff argues, better thought of in terms of relations between people, rather than distribution of things. But what are the appropriate social relations? Social egalitarians know what they are against – hierarchy, exploitation, domination, snobbery, servility, social exclusion – but find it much harder to say what they are for. Wolff suggested that social egalitarians ought to make a virtue of this difficulty and define social equality in terms of the avoidance of unequal social relations. A similar move is central to the way that we understand the possibility of theorizing an empirically informed theory of justice in Europe in ETHOS (see above), and indeed we named this approach 'Real World Political Philosophy' following Wolff. Instead of giving the ideal-theoretical articulations of a theory of justice a theoretically prior place (as proposed, for instance, by de Maagt), this approach starts with 'real world' investigation into *manifest injustices* before theorizing the appropriate responses to address those injustices using the tools of political philosophy.

While Wolff's keynote address proposed real world political philosophy as a method of pursuing empirically informed and problem-oriented justice research, it also problematized the foundational distinction between ideal and non-ideal theory. Four further contributions (by Jenkins, Landau, Crusmac and Teixeira) are clustered around different aspects of this taxonomic and conceptual question. Two aspects of this debate are especially relevant. First, as we have seen, the question of both the possibility and the status of the type of non-ideal theory that the ETHOS project pursues has been questioned in political philosophy. Giving a convincing response to detractors requires getting clear on the conceptual stakes. Second, we have noted that there are several, sometimes conflicting, notions of ideal and non-ideal theory in the literature, and have argued that not all are equally relevant for the ETHOS project. Clarity about where to draw the lines between these different notions and in which ways they overlap also requires some work of conceptual classification.

In the first section of this report we argued that the first 'fault line' of the ideal versus non-ideal debate in political philosophy was the least relevant to the ETHOS project. David Jenkins' arguments in his contribution to the workshop feed into developing this position. Jenkins rejects the standard approach to the conceptual distinction, which he took to be focused on ideal theory as perfect compliance, and non-ideal theory as imperfect compliance (this contribution is in Appendix 2). He argues that this way of theorizing non-ideal conditions rests not so much on a parsimonious description of injustice, but rather on an implicit and controversial understanding of the causes of injustice which fails to address a crucial distinction between *complicity* and *culpability*. Instead, Jenkins offers a more fine-grained analysis of the morally relevant ways in which agents can be responsible for injustice. Specifically, he argues that while complicity was both widespread and, in the real world, unavoidable, this should not distract us from the importance of identifying who is culpable for injustices.

De Maagt's paper, introduced above, argued in favour of a hierarchical structure of ideal over non-ideal theorizing. In contrast, against this and other claims for the priority of ideal theory or the dependence of non-ideal theory on ideal theory, Aaron Landau argues that a *systematic* non-ideal theory is possible and that, given our cognitive and epistemic limitations, the (ideal theoretical) prospect of achieving perfect justice cannot make a practical difference to pursuing the maximisation of expected justice (this contribution is in Appendix 2). The register of Landau's contribution is different than that of Jenkins, for instance, in that Landau takes as central the third of the three fault lines of the ideal versus non-ideal theory debate introduced above - that of end-state versus transitional theory. An end-state theory, recall, attempts to theorize perfect or utopian justice, while transitional theory is focused on improvements towards justice (or away from injustice). In that sense, the approach of Real World Political Philosophy developed here is very much in line with Landau's approach, although his is more radical than ours given his insistence, which echoes Amartya Sen's view, that end-state theory is largely redundant (for more detail on our contrasting positions on this point see both van den Brink's and Rippon and Zala's contributions in the appendix).

In the opening part of this report we discussed three fault-lines of the debate over ideal and non-ideal theory in political philosophy. Oana Crusmac's presentation had two main parts, the first of which was also primarily taxonomical and the second of which explores this typology for feminist and critical theory (the second part of this contribution is in Appendix 2). Crusmac divides the uses of non-ideal theory into two broad categories and five smaller subcategories. The first broad category follows the Rawlsian distinction between end-state (ideal) and transitional (non-ideal) theory, separating 'Rawlsian' non-ideal theory into three further blocks depending on the way in which transitions towards justice are theorized: in terms of short-term goals, in terms of long-term goals,

or in terms of realizing long-term goals through incremental short-term steps. All three of these ‘types’ of non-ideal theory share the characteristic of making non-ideal theory methodologically consequent to ideal theory (much like de Maagt does, but in contrast to, *inter alia*, Landau). The second broad category Crusmac identifies, in contrast, insists on the possible independence of non-ideal from ideal theorizing (this is the broad position we endorse in our methodology of Real World Political Philosophy - see above). She breaks this down into comparative non-ideal theory (following Sen’s ordinal approach) and independent non-ideal theory, which does use ideals and idealizations, but derives those in a fact-sensitive or ‘practice dependent’ fashion. The second part of Crusmac’s presentation sought to place certain feminist scholars within this conceptualization, particularly the last category of independent non-ideal theory (which Crusmac endorses). This leads Crusmac to develop two methodological approaches for pursuing independent non-ideal theory, the first she calls ‘abstraction without idealization’ and the second ‘group relations methodology’ (for more details see Appendix 2).

Pedro Teixeira, like many others in this literature (and indeed like this report) takes Rawls’ articulation of ideal and non-ideal theory as his starting point. Instead of working forwards from this theory, however, he worked backwards from it, attempting to explain *why* the discussion centred on the three poles introduced above (full versus partial compliance, the fact sensitivity or insensitivity of moral principles, and the orientation towards an end-state versus a transitional approach). Teixeira draws this genealogy back to the twin influences of Immanuel Kant and Hegel on Rawls’ theory of justice, and particularly to the attempt to combine a Kantian moral constructivism with insights drawn from Hegel’s idea of reconciliation. Teixeira does not stop at the genealogical task, however, but uses this background to argue - in the face of accounts we have discussed that postulate variously the priority of ideal over non-ideal theory (as in de Maagt) or, conversely, its independence (in Landau and Crusmac) - that ideal and non-ideal theory are two indispensable sides of any complete moral theory. This theoretical proximity and complementarity is stronger than Rippon and Zala’s notion of what we have referred to as the symbiosis of ideal and non-ideal theory above and, as such, poses a methodological challenge to ETHOS. Just as Teixeira may argue that ideal theorization without non-ideal considerations is fundamentally lacking, so he may question the possibility of pursuing non-ideal theory in the absence of an ideal-theoretical component.

The last thematic grouping of presentations gathers those that seek to apply a particular method of doing non-ideal theory to a political problem. Of course, these papers thus also concern methodology (like the first group) and the appropriate way to understand ideal and, especially, non-ideal theory (like the second). However, the primary purpose is more practice-oriented in the cases of the three presentations by Hickey, Kim and Benli. The methodological prescriptions and consequent applications of these three authors differ somewhat from the methodology of Real World Political Philosophy that we developed above. However, each tackles the transition from methodological prescription to recommendations in innovative ways. Indeed, we will see that our own application of the method we have developed to recommendations for the other ETHOS work packages, using the example of WP3’s work on disability rights (see the previous section), has been informed by various aspects of these approaches.

Using the example of climate justice, Colin Hickey develops an account of the individual duty to participate in the creation of just institutions. Instead of focusing centrally on establishing either a general or a more applied version of the duty to participate in the creation of just institutions, Hickey focuses on working out what such a duty most plausibly would consist in, if it obtains. He argues that the best way to understand such a duty is not to focus on the instrumental success (or not) in participating towards the creation of just institutions, but rather to

adopt the establishment of appropriately just institutions as personal ends. This shifts the goalposts of the debate on the appropriate level of idealization in people's compliance with justice demands, and what duties people have when others do not meet their duties (non-ideal theory - understood on the first of the three dimensions above). In its place, Hickey suggests that people's duties do not reflect a proportional 'slice' of the total effort needed to move from a situation of (climate) injustice to one of justice, but rather that each person independently has the duty to adopt contributing to the creation of just institutions as a personal end. While in the context of ETHOS we are not centrally concerned with the Rawlsian paradigm of full vs. partial compliance, Hickey's account gives an interesting example of precisely how to transcend this debate and theorize agents' duties of rectification differently (in this, Hickey's account is like Jenkins' account focused on the notion of culpability, discussed above).

Hwa Young Kim's paper engages a problem often overlooked in political philosophy, namely the various ways political philosophy can be used to best effect to address injustices practically (this contribution is in the appendix, below). This is a crucial part of the ETHOS project too, since, as Kim argues, even detailed and fact-sensitive or empirically informed theories of justice and fairness will have little impact if no one is motivated to act on their recommendations (in the ETHOS project structure this task is mainly pursued in work package 7). The context of Kim's argument lies in controversy over the degree to which variously idealized theories are likely to have an impact on people's motivations to pursue justice. David Miller, for example, holds that theories which demand too much (are 'utopian' or 'neo-Augustinian') are likely to be *demotivating* (2013), while G.A. Cohen thinks that precisely the radicalism and purity (in the sense of fact-insensitivity) of his theory of justice is more likely to galvanize people into action when compared to Rawls' theory which builds in, according to him, too many concessions to a lack of human motivation to meet the duties of justice perfectly (2008). In effect, the argument shows how the complexity and systemization of ideal theory undermines political conviction, and ought for that reason to be supplemented with other approaches to ensure justice can, in his words, be 'weaponised'.

Ali Emre Benli develops an account of non-ideal theory he termed a 'social justice approach', showing its application with reference to the question of refugee rescue missions in the Mediterranean. Benli's account is inspired by Sen's influential critique of ideal theory and his proposal of a comparative method. One of Benli's core insights is that pre- or proscriptive political theory (or 'action-guiding theory') is, in an important sense, unavoidably comparative. Choices about what to do in a particular situation are both bounded by that situation, and justified in comparison to other possible actions (or non-action). This leads Benli to articulate a standard to adjudicate what he calls 'accessible courses of action' in the face of the problems of complexity and moral disagreement. The second problem is particularly relevant for ETHOS - recall how we tackled the problem of reasonable and potentially incommensurable pluralism about value in the first section. In contrast to our solution to this difficulty, which focuses as a matter of moral urgency and epistemic contingency on 'manifest injustice', Benli proposes to look at the range of ethical positions that generate policy recommendations to an urgent injustice (the refugee crisis) and look for 'partial agreements' between the different perspectives to make policy recommendations.

We earlier indicated how disagreement about fundamental political or moral theories presents a challenge to applied political philosophy. It will be evident to readers of this summary of the workshop proceedings that the challenge is generalizable: if you put five philosophers in a room to discuss something, you will likely end up with at least six conflicting opinions. This is in part a consequence of the critical reasoning method of philosophy: since

philosophical views cannot typically be tested against empirical data, they must be tested theoretically, by rigorously assessing the best arguments that can be marshalled for and against them. This means that arguments and objections are prized, particularly objections to apparently well-established views and received wisdom, or arguments that seem compelling despite their defending outlandish ideas. In consequence, it is usually difficult to identify a 'philosophical consensus' about any substantial issue. We have rather sought, on the whole, to identify the state of the art by identifying the questions over which the current battles are fought, and summarizing the competing answers and the various arguments on each side.

Bibliography

- Altman, Barbara M. 2001. 'Disability definitions, models, classification schemes, and applications.' In *Handbook of Disability Studies*, ed. Gary L. Albrecht, Katherine D. Seelman, and Michael Bury: 97-122.
- Anderson, Elizabeth. 2010. *The Imperative of Integration*. Princeton: Princeton University Press.
- Berlin, Isaiah. 1969. 'Two concepts of liberty.' In *Four Essays on Liberty*: 118-172.
- Buchanan, Allen. 1996. 'Choosing who will be disabled: genetic intervention and the morality of inclusion.' *Social Philosophy and Policy* 13 (2): 159-85.
- Calder, Gideon. 2010. 'Pragmatism, Critical Theory and Democratic Inclusion,' *Ethics and Politics*, 12(1): 52-67.
- Cohen, G.A. 2003. 'Facts and principles.' *Philosophy & public affairs* 31(3): 211-245.
- Cohen, G.A. 2008. *Rescuing justice and equality*. Harvard University Press.
- Cole, Phillip and Christopher Heath Wellman. 2011. *Debating the Ethics of Immigration: Is There a Right to Exclude?* New York: Oxford University Press.
- Corvino, John and Maggie Gallagher. 2012. *Debating Same-Sex Marriage*. Oxford: Oxford University Press.
- Crowder, George. 2002. *Liberalism and value pluralism*. A&C Black.
- de Marneffe, Peter. 2010. *Liberalism and Prostitution*. New York: Oxford University Press.
- Dworkin, Ronald. 1981. 'What is Equality? Part 2: Equality of Resources.' *Philosophy and Public Affairs*, 10(4): 283-345.
- Erman, Eva, and Niklas Möller. 2016 'What distinguishes the practice-dependent approach to justice?' *Philosophy & Social Criticism* 42(1): 3-23.
- Fraser, Nancy. 1995. 'From Redistribution to Recognition: Dilemmas of Justice in a 'Post-Socialist' Age.' *New Left Law Review*, 212: 68-93.
- Fraser, Nancy. 2005. *Reframing Justice: The 2004 Spinoza Lectures*. Amsterdam: Van Gorcum.
- Fraser, Nancy, Hanne Marlene Dahl, Pauline Stoltz, and Rasmus Willig. 2004. 'Recognition, redistribution and representation in capitalist global society: An interview with Nancy Fraser.' *Acta Sociologica* 47(4): 374-382.
- Gliedman, John and William Roth. 1980. *The Unexpected Minority: Handicapped Children in America*, New York, NY: Harcourt Brace Jovanovich.
- Hamlin, Alan, and Zofia Stemplowska. 2012. 'Theory, ideal theory and the theory of ideals.' *Political Studies Review* 10.1: 48-62.
- Husak, Douglas and Peter de Marneffe. 2005. *The Legalization of Drugs*. Cambridge: Cambridge University Press.
- Jubb, Robert. 2009. 'Logical and epistemic foundationalism about grounding: The triviality of facts and principles.' *Res Publica* 15(4): 337-353.
- Levy, Jacob T. 2016. 'There Is No Such Thing as Ideal Theory.' *Social Philosophy and Policy* 33(1-2): 312-333.
- MacIntyre, Alasdair. 2013. *After virtue*. London: Bloomsbury Academic.

- Meckled-Garcia, Saladin. 2013. 'The practice-dependence red herring and better reasons for restricting the scope of justice.' *Raisons politiques* 3: 97-120.
- Miller, David. 2011. 'Taking up the Slack? Responsibility and Justice in Situations of Partial Compliance.' In *Responsibility and Distributive Justice*, eds. Zofia Stemplowska and Carl Knight. Oxford: Oxford University Press: 230-45.
- Miller, David. 2013. *Justice for earthlings: Essays in political philosophy*. Cambridge University Press.
- Mills, Charles W. 2017. *Black rights/white wrongs: the critique of racial liberalism*. Oxford University Press
- Pogge, Thomas. 2008. 'Cohen to the Rescue!' *Ratio* 21(4): 454-475.
- Rawls, John, 1999a. *A theory of justice: Revised edition*. Cambridge, MA: Harvard University Press.
- Rawls, John. 1999b. 'The Law of Peoples.' In *The Politics of Human Rights*, ed. Obrad Savic. London: Verso: 16-45.
- Rawls, John. 2003. 'Justice as Fairness.' In *Contemporary Political Theory: A Reader*, ed. Colin Farrelly. London: Sage: 13-21.
- Robeyns, Ingrid. 2012. 'Are transcendental theories of justice redundant?' *Journal of economic methodology* 19(2): 159-163.
- Ronzoni, Miriam, and Laura Valentini. 2008. 'On the meta-ethical status of constructivism: reflections on GA Cohen's "Facts and Principles"'. *Politics, Philosophy & Economics* 7(4): 403-422.
- Sangiovanni, Andrea. 2008. 'Justice and the Priority of Politics to Morality.' *Journal of political philosophy*, 16(2): 137-164.
- Sangiovanni, Andrea, 2016. 'How practices matter.' *Journal of Political Philosophy*, 24(1): 3-23.
- Scotch, Richard. K. and Kay Schriener. 1997. 'Disability as Human Variation: Implications for Policy.' *The Annals of the American Academy of Political and Social Science*, 549(1): 148-159.
- Sen, Amartya. 1990. 'More Than 100 Million Women Are Missing.' *The New York Review of Books*
- Sen, Amartya. 2006. 'What Do We Want from a Theory of Justice?' *The Journal of Philosophy* 103(5): 215-238.
- Sen, Amartya. 2011. *The Idea of Justice*. Cambridge, MA. Harvard University Press
- Shklar, Judith. 1990. *The Faces of Injustice*. Yale University Press
- Silvers, Anita. 1998. 'Formal Justice.' In *Disability, Difference, Discrimination: Perspectives on Justice in Bioethics and Public Policy*, eds. Anita Silvers, David Wasserman, and Mary B. Mahowald. Boston: Rowman & Littlefield: 13-146
- Simmons, A. John 2010. 'Ideal and Nonideal Theory.' *Philosophy & Public Affairs*, 38(1): 5-36.
- Stemplowska, Zofia, 2017. 'Non-ideal Theory.' *A Companion to Applied Philosophy*: 284-296.
- Stemplowska, Zofia and Adam Swift. 2014. 'Rawls on Ideal and Nonideal Theory.' In *A Companion to Rawls*, eds. J. Mandle and D. A. Reidy. Wiley: 112-127.
- Thomas, Carol. 1999. *Female Forms*. Buckingham, Philadelphia: Open University Press.
- UPIAS. 1976. *Fundamental Principles of Disability*. London: Union of the Physically Impaired Against Segregation.
- Valentini, Laura. 2011. 'Global Justice and Practice-Dependence: Conventionalism, Institutionalism, Functionalism.' *Journal of Political Philosophy* 19(4): 399-418.

- Valentini, Laura. 2012. 'Ideal vs. Non-ideal Theory: A Conceptual Map.' *Philosophy Compass*, 7(9): 654-664.
- Waldron, Jeremy. 1999. *Law and disagreement*. OUP, Oxford.
- Wasserman, David. 2014. 'Is Disability Discrimination Different?' In *Philosophical Foundations of Discrimination Law*, eds. Deborah Hellman and Sophia Moreau. Oxford: Oxford University Press: 269-277.
- Wasserman, David, Jerome Bickenbach, and Robert Wachbroit. 2005. 'Introduction.' In *Quality of Life and Human Difference: Genetic Testing, Health Care, and Disability*. New York: Cambridge University Press: 1-26.
- Wasserman, David, Adrienne Asch, Jeffrey Blustein and Daniel Putnam. 2013. 'Disability and Justice.' *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Available at: <http://plato.stanford.edu/archives/sum2013/entries/disability-justice>
- Wasserman, David, Adrienne Asch, Jeffrey Blustein, and Daniel Putnam. 2016. 'Disability: Definitions, Models, Experience.' *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Available at: <http://plato.stanford.edu/archives/win2015/entries/disability>
- Wiens, David. 2012. 'Prescribing institutions without ideal theory.' *Journal of Political Philosophy* 20(1): 45-70.
- Wolff, Jonathan. 2002. 'Addressing disadvantage and the human good.' *Journal of Applied Philosophy* 19 (3): 207-18.
- Wolff, Jonathan. 2009. 'Disability among equals.' In *Disability and Disadvantage*, eds. Kimberley Brownlee and Adam Cureton. Oxford: Oxford University Press: 112-137.
- Wolff, Jonathan. 2011. *Ethics and Public Policy: A Philosophical Inquiry*. Routledge.
- Wolff, Jonathan. 2015. 'Social equality and social inequality.' In *Social Equality: Essays on What It Means to be Equals*, eds. C Fourie and F Schuppert and I Wallimann-Helmer. Oxford University Press: 209-225
- Wolff, Jonathan, and Avner De-Shalit. 2013. 'On Fertile Functionings: A Response to Martha Nussbaum.' *Journal of Human Development and Capabilities* 14(1): 161-165.
- World Health Organization (WHO). 1980. *International Classification of Impairments, Disabilities and Handicaps: A Manual of Classification Relating to the Consequences of Disease*. Geneva.
- Young, Iris Marion. 2009. *Responsibility for Justice*. Oxford: Oxford University Press.

APPENDIX 1. Workshop Proceedings: Abstracts of Papers

Abstracts provided by authors from the *Ideal and Non-Ideal Theories of Justice* workshop, held on February 23-24, 2018 at Coimbra University, Coimbra, Portugal, as part of the ETHOS annual conference.

Abstracts are shown in the order in which they were presented at the workshop.

- ‘How Transcendental Arguments Can Be Used in (Applied) Normative Political Theory: An Alternative to Reflective Equilibrium’, Sem de Maagt (Utrecht University)

Although many people working in the more applied areas of ethics might be willing to admit that the method of reflective equilibrium is not a perfect method of moral justification, there is at the same time a widespread conviction that there simply is no practicable alternative. The aim of this paper is to show how one such an alternative, transcendental argumentation, could be put to work in applied normative political theory. In this way, I hope to show that there is in fact an alternative, practicable method of moral justification in applied ethics.

- ‘Non-Ideal Theory and the Duty to Participate in the Creation of Just Institutions’, Colin Hickey (Utrecht University)

In this paper, I develop an account of the individual duty to participate in the creation of just institutions. Many scholars have gestured at such duties, but surprisingly rarely are they critically explored in depth. This is unfortunate, given the significant role such duties have at the heart of discussions of ideal/non-ideal theory, and with respect to understanding some of our most crucial practical duties in contemporary moral and political contexts.

My main aim isn't to argue for the existence of either a general duty to participate in the creation of just institutions or any specific duty for particular institutions, per se (though I begin with a range of considerations regarding such arguments, and will employ the case of climate change and the creation of just climate governing

institutions as an important model throughout). Instead, assuming some such duties have been plausibly established, I use this opportunity to explore some important theoretical considerations about such duties.

I suggest that the duty to participate in the creation of just institutions is most essentially a duty to adopt a certain kind of end. There are often many routes toward demonstrating that one has sufficiently adopted that end, and what would demonstrate such adoption depends on historical, social, and material facts of individuals in question. However, I suggest some resources for thinking through a floor past which one must certainly cross in order to plausibly count as having adopted the end and a ceiling on what could plausibly be required of one to count as having adopted the end. Moreover, I suggest some important general metrics by which we should begin to characterize individuals, in order to understand the profile of their duty and have the methodological resources for interpreting its satisfaction conditions.

Finally, I suggest that in many cases we shouldn't think of each individual's duty to participate in the creation of just institutions as some finite percentage of the total agential effort required for their creation, like apportioning 'wedges of a pie'. Instead, I maintain, if everyone actually did what, for them, would constitute sufficiently adopting the end, often it would amount to more total agential activity than would have been strictly required to create the institutions. That is, the overall system of duties often should have built in redundancy, which tells us something informative about how others' noncompliance may/may not affect the shape, strength, and content of one's own duty.

- 'The Importance of Ideal and Non-Ideal Theory', Simon Rippon & Miklós Zala (Central European University)

Ideal theory in political philosophy is theory of justice that idealizes in one or more ways. It may unrealistically assume full compliance, ignore empirical complications, envisage utopias we could not possibly get to from here, or make demands that people cannot be expected to live up to.

Many critics of ideal theory argue that it does not provide useful guidance for what to do in the real world here and now, some argue that ideal theorising blinds us to certain kinds of real-world structural injustice, and some argue that it is a methodologically dubious or altogether incoherent. These critics may endorse a non-ideal theoretic approach, starting from observations of actual empirical circumstances and injustices and normatively examining potential improvements.

In this paper, we argue that there is no serious case to be made for failing to engage in non-ideal theory, since we want political theory to help solve the very practical problem of what to do; a problem that ideal theory cannot directly address. This means that political philosophy can and must engage with the empirical social sciences, and with various normative problems that simply disappear at the level of (completely) ideal theory. But this is not to negate the importance of ideal theory. Many doubts about ideal theory arise from misunderstandings of what it is and what it is for. Applying or aiming at realizing the principles of an ideal theory, here and now, may well be inadvisable, but ideal theory is not designed as a procedure for deciding what to do here and now. Pointing out that something is bad at doing what it may be misused to do is not a legitimate criticism of it. We argue, moreover, that non-ideal theorizing implicitly appeals to at least an inchoate conception of the good, a conception that ideal theory can be used to refine and clarify. This is because ideal theory is not best understood simply as a description of a perfect utopia, but rather as a system of principles and/or values together with reasons for accepting them. The grasp of such a theory does not give us a mere picture of the perfect, but rather a deeper understanding of the good. Since this in turn may yield a clearer understanding of what makes real-world alternatives better or worse, it turns out that ideal theory can after all play a useful role, if what we want to do is make the world we live in less unjust.

- ‘Ideal and Systematic Theory’, Aaron Landau (Columbia University)

The practical motivation for constructing an “ideal theory” of a perfectly just society is that it is essential for justifying some of the principles for “nonideal theory,” which tells us how to act when there is substantial injustice. As John Rawls says, “[t]he reason for beginning with ideal theory is that it provides ... the only basis for a systematic grasp of [the] more pressing problems” of nonideal theory (1999a p.8). Among the various arguments for this “priority” or “dependence thesis,” two are especially prominent. First, ideal theory is necessary for establishing priority rules (e.g. Rawls’s famous lexical ordering) that show us which injustices to remove before others. Second, we must do ideal theory because we need to know how we should transition to a fully just society. In this paper, I argue that these claims are mistaken and that we can construct a systematic nonideal theory without knowledge of the ideal.

I begin by carefully defining “ideal theory,” “ideal principles,” “nonideal theory,” and the “dependence thesis” in order to set the terms of the debate. Importantly, I distinguish ideal theory and ideal principles from other kinds of moral theory and principles that may be independently available to nonideal theory. The dependence

claim is true if and only if ideal theory can make a contribution to nonideal theory that these other types of theory and principles cannot. I show that we can establish both principles of justice and priority rules for non-ideal situations apart from ideal theory not only for consequentialism but also contractualism, for which equality and fairness are basic values. Moreover, I argue that a structurally identical critique applies to many moral frameworks, including Rawls's justice as fairness and political liberalism. Finally, I examine the influential "realistic utopia thesis" that ideal theory is necessary because we need to know how to transition to our ultimate goal of a just society. I propose that the argument for this thesis is best interpreted as being based on a moral requirement to seek the greatest expected justice, other things being equal. However, given our epistemic and cognitive limitations, the prospect of achieving perfect justice cannot make any difference to the maximization of expected justice. It follows that, in our contemporary nonideal circumstances, we can fulfill our duty to maximize expected justice without ideal theory.

- Keynote lecture: 'Real World Egalitarianism', Jonathan Wolff (University of Oxford)

Professor Jonathan Wolff opened his keynote lecture by focussing on the semantic distinction between 'ideal' and 'non-ideal' theory. He argued ought to be rejected, as it makes latter appear parasitic on former. Wolff proposed that the term 'real world' could be used to replace 'non-ideal'.

Contemporary egalitarianism dominated in 1980s and 1990s by what Anderson came to call 'luck egalitarianism', emphasising the distinction between circumstance and choice, and supposing that justice requires the equalization of circumstances, but allows people to reap benefits and burdens of freely-made choices. 'Luck egalitarianism has two key features: it is distributive and responsibility-sensitive. Both of these, Wolff argued, are given too much emphasis. He discussed his 1998 paper (Fairness, Respect and the Egalitarian Ethos) which detailed his reasons for rejecting both. Responsibility-sensitivity requires intrusive mechanism that are inconsistent with the idea of a society of equals.

The notion of a society of equals is, Wolff argued, better thought of in terms of relations between people, rather than distribution of things. But what are the appropriate social relations? Social egalitarians know what they are against – hierarchy, exploitation, domination, snobbery, servility, social exclusion – but find it much harder to say what they are for. Wolff suggested that *social egalitarians* ought to make a virtue of this difficulty and define social equality in terms of the avoidance of unequal social relations.

Poverty, Wolff argued, is very interesting in this connection. Poverty is largely overlooked by ideal theory, as it is not part of any plausible theory of 'ideal' or maximal justice; in the real world, however, poverty is very important. At a first glance, however, dealing with poverty from the perspective of justice appears to be possible with a distributive, responsibility-sensitive theory like luck egalitarianism: poverty concerns the possession of resources, and such theories seem to allow one to articulate a principled cut-off between those there is a duty to help and those for whom there isn't. However, this approach to resources in the context of poverty ignores a key notion in contemporary scholarship on poverty – not having enough to fit in – which is a relational notion (non-exclusion), operationalised in resource terms, rather than purely a matter of resources.

Good social policy prevents poverty, and provides collective goods that all can enjoy whether in poverty or not (schools, hospitals, transport, public realm etc). Some targeting inevitable but much can be avoided. Wolff argued that it is not always necessary to identify people as being in poverty in order to help them overcome poverty. In conclusion, real world egalitarianism starts from the problems we encounter in real world, linking where possible with social movements, and using philosophical reasoning to conceptualise problems and co-produce solutions, rather than setting out ideal constitution.

- 'Ideal and Nonideal Theory: A Conceptual Analysis of the Divide', Oana Crusmac (University of Bucharest)

The presentation aims to present a detailed analysis of the ideal/nonideal debate and to investigate which outcome of this debate better suits a gender-sensitive theory of justice. I will present the Rawlsian distinction between ideal and nonideal theory as well as the classical dispute between ideal and nonideal theorists. I will also analyse whether and how ideal and nonideal theory can be interrelated. Here I will try to distinguish between five subtypes of nonideal theory and I will argue that only two of them manage to meet the basic criteria in order to be taken into account in a sound theory of distributive justice. Given that since its beginning as a term (in Rawls' *A Theory of Justice*), ideal theory has been used in the context of domestic distributive justice and most of the critiques followed this direction, in this paper I will limit my analysis to domestic justice in order to be able to incorporate some feminist recommendations.

Within the vast academic work on the methodology of ideal versus nonideal theory, many feminist and critical race theorists tend to form a common front against universal and abstract norms promoted by canonical political theory and, more notably, by abstract liberalism. Another common perspective shared by feminists and

race theorists is their critique and, most of the time, rejection of the distributive paradigm in political thought as the sole and universal solution to ending injustice and inequality.

The presentation will be split into three main parts, according to the three main points of departure specific to the ideal–nonideal debate: disputes over the structure and characteristics of both ideal and nonideal theory, and disputes over the relationship between the two. The first section will address the main characteristics and criticism targeting Rawls' ideal theory. The three main criticisms against ideal theory based on the rejection of one of Rawls' three idealized assumptions, namely: strict compliance, favourable circumstances, and idealized, rational agents. The second section will pursue the same task but concerning nonideal theory. Here I will distinguish between two main types of nonideal theory (Rawlsian or transitional nonideal theory, and independent nonideal theory), each having several subtypes. I will argue that three of these five subtypes are not accurate depictions of what nonideal theory is or should be, and that only two subtypes represent the most important candidates on the conception of nonideal theory. The third section focuses on the core of the ideal–nonideal divide [*this section constitutes Crusmac's draft paper included in the following section*], namely the relationship between the two, taking into account the five subtypes of nonideal theory and the charges raised by feminist theorists against the manner in which abstraction is employed in ideal theory.

- 'Weaponising Justice and Equality', Hwa Young Kim (University of Warwick)

How do political philosophers convert their theories of justice to practice? Most political philosophers follow the path set out by John Rawls. To put it simplistically, we start off with an ideal theory. It not only assigns rights and duties and a fair division of social advantages, it also provides an 'ultimate target'. We then move on to non-ideal theory. By taking a more empirically-sensitive approach to non-ideal/real world circumstances, non-ideal theory can provide a road map towards the ultimate target provided by ideal theory.

I believe there is something missing with this model of theory and practice. An incredibly detailed and empirically-sensitive road map will not have any significant impact on social change and progress if no one reads (or cares about) the map. The roadmap is important but there needs to be some aspect of the theory (both ideal and non-ideal) that inspires people to fight for that particular conception of justice in the political arena.

G. A. Cohen provides another account of the relationship between theory and practice that provides a solution to this problem. Unlike Rawls, Cohen excludes considerations of feasibility, publicity, and stability from justice and other fundamental normative principles. This would suggest that Cohen would be the last person to

turn to when thinking about political practicality. However, one of the reasons why Cohen excludes these features from the concept of justice is that he holds a different model of theory and practice. Ideal theory is not meant to provide a blueprint of an ideal world. Instead, Cohen argues that theory provides a ‘depth of conviction’ necessary for political agents to use the theory ‘as a weapon’ in the endless struggle of politics (Cohen, 2011, 212). I refer to this as the convictional role of justice.

Can Rawlsian/Constructivist theories of justice play this conviction role? I believe it can. However, there are features of the constructivist/contractualist methodology that limit its ability to play this role, such as its complexity and its aim to develop consensus. Because of this, I argue that we need a more complex model of theory and practice than the one provided by Rawls. To put it simplistically, we should follow the Rawlsian model to provide a clear roadmap and to generate consensus among reasonable people. We should follow the Cohenite model to inspire reasonable people against unreasonable people in the endless struggle of politics.

- ‘Keeping Non-Ideal Theory Real’, David Jenkins (University of Warwick)

Contemporary political theory is often criticised for being too far removed from the realities of political life and thus unable to tackle the truly pressing, collective problems of our day. The move toward theory that is capable of tackling these problems, begins with first getting clear on the reality of our shared situation. One common approach to gaining this clarity has been to define non-ideal conditions in terms of deviation from an ideal assumption of ‘full compliance’: When theorising about justice within ideal conditions there is an assumption that the relevant individuals are taken to both know and do what is expected of them by the theory of justice, and to be fully assured that others will also act according to those same demands. Partial compliance is thus taken to be a failure to realise this ideal condition: People fail to conform to the demands made of them under ideal conditions, and the assurance of others’ compliance is also missing. This is intended to describe, in a general way, the conditions of the real world.

In this paper, I argue that this way of theorising about non-ideal conditions is deeply mistaken, that it rests not so much on a parsimonious description of injustice, but rather on an implicit and controversial understanding of the causes of injustice, rooting them in a blanket failure of all people to do what is expected of them. This fails, first and foremost, to address a crucial distinction between complicity and culpability. By bringing these two concepts to the front and centre of our theorising about non-ideal conditions, we are better able to first recognise the ubiquity and inescapability of complicity, something which does a better job than ‘partial compliance’ in

accurately describing a position we do all share, whilst also distinguishing this condition from the powers and actions that render only certain agents culpable as well as complicit. By so doing, theorists achieve a better sense of both the perplexities and challenges presented by contemporary political reality, while also pointing to possible resources by which the perpetrators of contemporary injustice might be identified and confronted. I conclude with a description of the necessarily subversive and disruptive work that needs to be engaged with if the properly-motivated complicit are to effectively resist the powerful and culpable, without also then succumbing to the charge that they themselves become culpable.

- 'Rawls, Kant and Hegel: The Problem of the Ideal and Non-Ideal Division', Pedro Teixeira (Free University of Berlin)

The purpose of my PhD project is to contribute to the discussion on the limits and possibilities of the ideal and non-ideal theory distinction in political philosophy. I hope to accomplish this by looking into two moments in German philosophy that offer substantially different accounts of similar problems: first, the development of Hegel's philosophy until and including Marx, inasmuch as it revealed the incompatible views between right and left Hegelians regarding the embodiment of values; and second, in the current split between the political theories of Jürgen Habermas and Axel Honneth, that originates in their different interpretations of the Kantian and the Hegelian projects.

In this presentation I will attempt a reconstruction of the ideal/non-ideal step by looking at the twin influence of Kant and Hegel on Rawls' work. I will task myself with showing that the specific manner in which he merged some of his influences (here, Kant and Hegel) explains the need for introducing the methodological subdivisions of ideal and non-ideal, as well as accounting for a certain hesitation on how to characterize them. In particular, I will argue that Rawls' efforts to combine a political form of Kant's moral constructivism with a Hegel-inspired notion of reconciliation may have led him to characterise the ideal and non-ideal distinction into three distinct levels, which may account for the contemporary discussion revolving around three main axes (cf. Valentini, 2012): stability (full vs. partial compliance), feasibility (utopia vs. realism) and orientation (end-state vs. transition). I will try to show that the division of his theoretical enterprise into two parts was a necessary but problematic turn, from Rawls' own standpoint. I will conclude by suggesting the unification of the ideal and non-ideal as two necessary moments of any complete social theory.

- ‘A Social Judgment Approach’, Ali Emre Benli (UC Louvain)

An action-guiding political theory needs to tell us what should be done in the here and now. It is essentially comparative, as it needs to delineate moral advantages of accessible courses of action. It should also be determinate enough to deliver in the face of actual problems. This is not an easy task. We are ultimately faced with two major challenges. First, given the complexity of actual circumstances, it is hard to determine the scope and character of empirical information relevant for a case under investigation. Second, given extensive disagreement on moral issues, it is hard to specify duties without imposing one’s own moral view on others. In this article, I develop what I call a social judgment approach to meet this task and demonstrate its merits by working through a case of European Union policy regarding search and rescue missions in the Mediterranean.

The approach, inspired by the work of Amartya Sen in *The Idea of Justice*, focuses on a particular problem and compares accessible courses of action in terms of justice. The extent of the problem and availability of remedies allow us to identify moral agreements on what should be done without the need to work out what ideal justice would involve. In actual circumstances, there is at play a whole range of ethical perspectives that govern our judgments and each individual tends to have a preconceived idea on the relative importance of each perspective. Yet, two assumptions allow us to identify significant agreements. First, so far as ethical perspectives are impartial, all individuals should give some value to each perspective. Second, relative valuations of ethical perspectives transform in light of empirical information regarding particular contexts.

For example, although the set of moral values that ground a restrictive refugee regime and the set that grounds a right-oriented refugee regime will diverge, they may nevertheless both justify extensive rescue missions in the Mediterranean. Partial agreements relative to the context point at courses of action that lead to less unjust social states. The outcome of the evaluation is a moral judgment that can always be improved in light of new empirical information and critical moral reasoning. Yet, it can still ground legitimate duties, as the judgment is not an imposition of one’s own moral view, but grounded on an agreement of all parties.

- ‘Ideal theory and political perspectivism’, Bert van den Brink (Utrecht University/ University College Roosevelt)

Since political practices in liberal-democratic societies are full of different and conflicting political perspectives, which are guided by their own and sometimes conflicting ideals, it is unlikely that through philosophical reflection on these perspectives, we will arrive at just one ideal theory of political cooperation. In my talk, I will ask whether this means that we can do without ideal theory in political philosophy. I will argue that we cannot, and that even a perspectival approach to political theory cannot do without ideal-theoretical approaches. I will discuss politically liberal, deliberatively democratic, and more agnostic approaches to democracy.

APPENDIX 2: Selected Papers from the Workshop

Overview

This appendix includes six of the working papers from the *Ideal and Non-Ideal Theories of Justice* workshop, held on February 23-24, 2018 at Coimbra University, Coimbra, Portugal, as part of the ETHOS annual conference. The working papers of all those participants who gave consent are included here in draft form, without prejudice to potential future publication of final versions. (Several participants requested that their work not be included in here beyond their abstract, or that only parts of their presentations were included, either to guarantee no complications with future publication or because the work was not in a format that they wanted to see distributed beyond the workshop.). The selected papers are:

- Van den Brink, Bert, 'Ideal Theory as Politics' p.1-10
- Rippon, Simon and Miklós Zala, 'The Importance of Ideal and Non-Ideal Theory', 11-28
- Crusmac, Oana, 'Feminist political theory and non-ideal theory as its methodology', 29-42
- Jenkins, David, 'Keeping Non-Ideal Theory Real', 43-67
- Kim, Hwa Young, 'Weaponising Justice and Equality', 68-87
- Landau, Aaron, 'Ideal and Systematic Theory', 88-96

In contrast to the main body of this report, the papers included here reflect the views only of the named authors.

The first two working papers are by certain members of the ETHOS WP2 team and constitute their own draft formulations of ongoing research in the context of the ETHOS project's engagement with bridging empirical and normative theories of justice. The latter four papers are unedited working papers, and are not the product of ETHOS researchers.

These are draft papers. Please do not cite without prior permission from the author(s).

1. Ideal Theory as Politics

WORKING PAPER by Bert van den Brink (University College Roosevelt/ Utrecht University)

In this sketch of a future article, I aim to bring together two things that are not usually brought together. On the one hand, I will argue that normative political theories such as theories of justice cannot avoid being embedded in the societies of their time and to play a political role in them (sections 1 and 2). On the other hand, I will argue that there is no conflict whatsoever in acknowledging the contextual rootedness of these theories on the one hand, and gaining radically new understandings of what these contexts allow of through ideal-theoretical or idealizing thinking on the other (section 3). I will end my sketch with some reflections on what my argument may imply for our understanding of the activity of political philosophy.

§.1 - Learning from the history of political philosophy

Political philosophers can learn a lot from the history of their trade. And this not just because it is important for the trained political philosopher to know what great political philosophers such as Plato, Aristotle, Machiavelli, Hobbes, Locke, Mill, Arendt, and Rawls thought. Even more important is a firm acknowledgement that none of these thinkers was in the business of spelling out eternal general principles and rules for every possible society. Every philosopher will grapple with abstract conceptual issues that may seem quite timeless. But thinking of these philosophers as setting up specific political theories concerned with specific political questions of their time makes us see that their work is much deeper ingrained in the political vocabularies of their times than is often acknowledged.

One of the texts in the history of political thought that has brought this insight home to me is Quentin Skinner's *Liberty before Liberalism*.¹ Skinner articulates an epiphany he had as an undergraduate student: "[d]uring my second year as an undergraduate student, Peter Laslett published his definitive edition of John Locke's *Two Treatises of Government*. From Laslett's Introduction I learnt that, while there is perhaps no harm in thinking of Locke's *Two Treatises* as a classic defence of contractarianism, you will have no hope of understanding his text unless you recognize that it was primarily intended as an intervention in a particular crisis of English royalism under Charles II, and was written from an identifiable position on the spectrum of political debate in the early 1680s."² What strikes me most here is the observation that a famous political theory, which is indeed best known as a classic defence of one of the most dominant approaches to normative political theory, was intended as an intervention in a particular political constellation.

Skinner claims that one way to think of the relation between politics and political philosophy is to acknowledge that, "what it is possible to do in politics is generally limited by what it is possible to legitimize. What you can hope to legitimize, however, depends on what courses of action you can plausibly range under existing normative principles."³ Seen in this way, the primacy in politics is always with practice. Practice, however, is not a chaotic vat of opinion and desire; it rather is structured by authoritative principles and institutional orders, which have their own legitimations. Articulating such legitimations in the most systematic manner possible is the work of philosophers, who take positions on political spectrums, but phrase legitimations of "what is possible" in narratives that seek to be accepted by all as rational and reasonable.

Once we acknowledge that the great political theories are not time-transcending manifestos of philosophical truth, but rather classic statements of positions in long gone by political debates, we cannot force them to guide us from a viewpoint *sub specie aeternitatis*. That means, among many other things, that when

¹ Quentin Skinner, *Liberty before Liberalism*, Cambridge: Cambridge University Press, 1998 (Canto Classics digital).

² Ibid, loc. 1654 Canto Classics digital (Chapter 3).

³ Ibid, loc. 1682

faced with, for instance, questions of political authority, individual freedom, and religious toleration, we cannot simply ‘apply’ our Lockean theories of political authority and religious toleration to our current political and societal constellations.

And yet, our philosophical and political practices bear traces of that applied approach to great philosophy. Liberal authors draw lines from John Locke to Adam Smith, Immanuel Kant, John Stuart Mill, T.H. Green, Isaiah Berlin, and John Rawls to sketch the contours of the venerable tradition of thought that they stand in. Life, individual liberty, self-ownership, market, possession, and religious toleration are some of the seemingly time-transcending values that are supported by these philosophers. What could possibly be wrong with working with a framework inferred from the work of these great minds, even if they were politically positioned in their own times?

In a much-quoted passage Skinner answers this question as follows: “[I]t is remarkably difficult to avoid falling under the spell of our own intellectual heritage. As we analyse and reflect on our normative concepts, it is easy to become bewitched into believing that the ways of thinking about them bequeathed to us by the mainstream of our intellectual traditions must be *the* ways of thinking about them.”⁴ To avoid falling under this spell is remaining aware that one’s legitimizing framework for political (theoretical) judgment is not given *sub specie aeternitatis*, but rather the result of a choice for certain positions that have been taken on political spectrums from one generation to the next. This insight opens up new possibilities, indeed, it “is there to prevent us from becoming too readily bewitched” [...] by “the values embodied in our present way of life”.⁵

The substantive point of Skinner’s essay is to bring this methodological insight to philosophical debates about liberty. Until the 1990, these debates had for decades circled around the seemingly inescapable binary opposition of (liberal) ‘negative’ liberty and (conservative and authoritarian) ‘positive’ liberty. In his essay, Skinner brings to the fore (‘excavates’, is his own term) an alternative ‘neo-roman’ or republican theory of liberty, with

⁴ Ibid., loc. 1775.

⁵ Ibid.

historical roots in Roman legal thought, Machiavelli, and James Harrington, and with a contemporary protagonists in Philip Pettit and, indeed, Skinner himself.⁶ This intellectual tradition agrees with the liberal tradition that the liberty of individual citizens should always be respected; a stance that doctrines of positive liberty do not unqualifiedly acknowledge. Yet, the neo-Roman theory pledges that the liberal focus on the mere absence of illegitimate interference in the private sphere is not a sufficient guarantee against the wrong of harming individual liberty. It rather focuses on the need to organize state power in such a way that arbitrary and illegitimate interference in the private sphere cannot occur.

Coming back to the accusation of antiquarianism directed at Skinner, it now becomes clear that although this understanding of the history of political philosophy does not lend us direct tools for applying ‘eternal’ philosophical truths to practice, it does have the power to enable us to step back and consider options. The philosophical point and normative impact of considering various intellectual traditions and philosophical theories is that we can open up our inherited intellectual and political horizon and improve the quality of our own reasoning and eventually, our own judgment in politics and in political theory.

§.2 – From the history of political philosophy to political philosophy in the present

Let us bring Skinner’s insight to some of the most influential normative political theories of our own time. This would make us look at John Rawls’ politically liberal conception of justice and Jürgen Habermas’ theory of deliberative democracy as normative accounts of politics that *can be legitimized* under existing normative principles and that were formulated, however loosely, from an identifiable political position in their societies. I will argue that this is 1) an entirely plausible reading of the work of these authors and 2) that it does not in any way lessen the normative appeal of their theories.

⁶ Ibid.

1) At first sight, the reader's reaction may be that, as normative political theories, political liberalism and deliberative democracy aim to be more than theories that *can be legitimized* under an existing normative order. Some may want to argue that we need these theories to probe the legitimacy of the existing normative order at stake, not as illustrations of what can be legitimized under the existing order. But what can be legitimized under the existing order is not necessarily *just* the existing order. For instance, when it comes to its conceptualization and normative account of justice, political liberalism claims to be the *most reasonable* political theory for its day and age (Rawls indeed works with that restriction) and sketches a framework that can both be legitimized – i.e. made acceptable as rational and reasonable – and articulates a conception of political justice that is both intelligible to most and a critical yardstick. It most certainly cannot be understood as a blind affirmation of an existing normative order.

An even stronger claim characterizes the theory of deliberative democracy. Although it starts from quite widely accepted and uncontroversial assumptions as to the fundamental dignity, equality and mutual dependence of all members of society, it claims to offer a philosophical reconstruction and legitimation of *necessary normative presuppositions* of agency that everyone who acts under the institutions of democracy under the rule of law has to accept. Again, what emerges is a normative theory that in many ways fits well with the existing normative order in democratic societies under the rule of law, but can be used as a critical yardstick for evaluating the actual workings of such. Both of these theories argue that, all things considered, we are bound to accept them by virtue of our being rational and reasonable, and act from them – or in their spirit – in (part of) our political interactions.

Looking at the normative theories of political liberalism and deliberative democracy through Skinner's lens we have to assume that political theories tend to be developed "from an identifiable position on the spectrum of political debate." I don't think that this is meant to say that political theories tend to have a clearly identifiable political 'colour' in their societies of origin. I rather take this to imply that the basic assumptions and the normative outlook characteristic of such theories tend to be such that they take a position on some of the most important conceptual and normative questions of their times.

Habermas' work is rooted in the sociological theory of modernization and rationalization.⁷ He argues that in fully rationalized modern life worlds, no claim to authority of, for instance, a religious, traditional, political, or academic kind is immune to what he has labelled the "unforced force of the better argument."⁸ Deliberative democracy is born from a post-traditional, anti-authoritarian, egalitarian, and procedural take on democratic politics. It works with a decentralized account of the ultimate source of legitimate political power, according to which legitimate power claims in politics are those that will be accepted after a fairly structured deliberation in which the *truth* of statements about objective states of affairs, the *correctness* of statements about social norms, morality, and legislation, and the *truthfulness* of communicative agents may be tested.⁹ What exactly is a fairly structured democratic deliberation is a very complex question, which many theorists have worked on. I cannot go into that question here. For the purpose of my argument, the most important thing is to note that the basic choice for a post-traditional, anti-authoritarian, egalitarian and procedural approach makes the theory of deliberative democracy take an identifiable stance on some of the great conceptual and normative issues in politics of the late-twentieth and early twenty-first Century.

For political liberalism, the reconstruction of an identifiable position on the political spectrum is to be approached a bit differently. Most clearly in his 1993 book, *Political Liberalism* and in his 1999 article "The Idea of Public Reason Revisited,"¹⁰ Rawls has stated that his normative theory of political justice is to be understood as a systematic reflection on the potential of an existing political tradition with respect to dealing with questions of socio-economic distribution and political representation in the face of persistent societal pluralism, i.e. a variety of personal moral, philosophical and religious beliefs about the good, and hence a variety of conceptions of

⁷ Jürgen Habermas, *Theory of Communicative Action*, 2 volumes. Translated by Thomas A. McCarthy. Boston, Mass.: Beacon Press, 1984 and 1987.

⁸ Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Translated by William Rehg. Cambridge, MA: MIT Press, 1996, p. 306.

⁹ Habermas, *Communicative Action*, vol. 1.

¹⁰ John Rawls, *Political Liberalism*, New York: Columbia University Press, 1993. "The Idea of Public Reason Revisited," *University of Chicago Law Review*, 64/3 (1997): 765-807.

desert, dignity, and ultimate life goals. The broadly shared way of dealing with such questions of justice he labels liberalism, and his particular interpretation of it he labels ‘political liberalism’. This liberalism is ‘political’ because it deliberately stays on the surface of deep societal pluralism, and proposes that, when engaged in political debate, citizens active within the basic structure of the political system (its main political, legal and economic institutions) should not claim the truth or correctness of their personal “comprehensive” conceptions of desert, dignity and goals in life, but work with reasons and commonly held political goods shared by those who think of themselves as free and equal *citizens*, who engage in a mutually beneficial form of social cooperation. Although this brand of political liberalism does not claim a specific political ‘colour’, it clearly takes some far-going conceptual decisions with important normative implications that cannot possibly result in a neutral position on the political spectrum in Rawls’ day and age.

2) I already said why the rootedness of these theories in their times does not lessen their normative appeal. There most certainly is no perfect fit between these normative theories and the workings of the concrete political and legal institutions in society. They are rather rigorously reasoned accounts of how we can understand ourselves as co-authors of these institutions. The ideas central to these accounts have a considerable grip on the philosophical and political imagination with respect to the institutions of the liberal-democratic state and many of its citizens. They find uptake because they fit normative vocabularies with a much wider appeal in society, and these normative vocabularies in their turn have been more systematically thought-through with the help of these theories.

Further than that, and as a consolation for those who are sceptical of understanding political philosophy as a politically embedded activity, the more abstract normative appeal of these theories is not necessarily being pulled down by the understanding of their rootedness that I proposed. Indeed, it may very well be that some theorists of justice accept elements of the deliberative democracy because of the linguistic and logical strengths of Habermas’ formal-pragmatic reconstruction of communicative rationality. Similarly, many have been convinced by political liberalism not because of their adherence to a particular political tradition, but through the moral appeal and systematic rigor of Rawls’s methodological device of the original position.

§.3 – Bringing in ideal theory: the strength of the as if

Ideal theory and counterfactual idealizations often play an important role in normative political theory. They most certainly do in the theoretical frameworks that I have been discussing in the previous section. I want to claim that

there is there is no conflict whatsoever in acknowledging the contextual rootedness of these theories on the one hand, and gaining radically new understandings of what these contexts allow of through ideal-theoretical or idealizing thinking on the other.

John Rawls is well-known for using ideal theory as a means to articulate moral ideals and obligations that we may not necessarily recognize from our everyday, ‘rooted’ intuitions and beliefs. He starts his reflections on justice by working under idealizing assumptions, which he lifts once more concrete and non-ideal theoretical questions of justice have to be answered. The most important idealizing assumptions concern, first, the compliance of all in the relevant society with justified principles of justice and, second, relatively favourable social conditions that make it feasible to abide by such principles. The idea is that by working under idealizing assumptions, we can get a clear picture of which principles of justice we would accept under conditions that can be wholeheartedly trusted by all addressed. According to Rawls, this helps us formulate a ‘realistic utopian’ picture of what we may hope for with regard to our social order and which guidance we may seek in answering non-ideal questions of justice in the real world.¹¹

That question of guidance is a much-debated issue. Can we know what to do in what we deem unjust situations when we do not have a positive ideal theory of justice as a reference point?¹² Instead of engaging with that debate about which ideal or non-ideal approach in political philosophy we *must* choose, for my purposes it is more interesting to ask whether the account of the rootedness of political philosophy that I started with can be combined with an ideal theoretical approach.

Rawls’ theory of political liberalism certainly shows us that this is possible. His theory helps us imagine a picture of justice that we may accept *if* we think of ourselves as citizens in a mutually beneficial form of social cooperation, who think of themselves as free, equal, rational and reasonable, and confronted with a reasonable

¹¹ For a brief but very clear positioning of the importance of ideal theory that I found helpful, see Leif Wenar, “John Rawls,” *The Stanford Encyclopedia of Philosophy*, (Spring 2017 Edition), Edward N. Zalta (ed.).

¹² See for an influential ‘non ideal’ criticism of Rawls’ approach Amartya Sen, *The Idea of Justice*, London: Penguin, 2010.

form of pluralism. A lot of *ifs*, you may say. Rawlsian ideal theory is perhaps best understood as a theory of the *as if*. It asks us to look at ourselves *as if* we were all this – and then to answer the question whether we could accept these terms of justice, provided that all others would and social conditions would be sufficiently favourable. If our answer is yes then we have bound ourselves to a hypothetical situation, not to a real one. But by doing so, we acknowledge that there is an important, possibly universalizable point to this theory of justice. The ideal may subsequently help us work towards favourable conditions for and greater compliance with this particular theory of justice.

The philosophy of the *as if* allows us to invite both all those who are situated at a similar position on the political spectrum of our time and those who are differently situated to follow our thought and vision for a while, without being bound by anything. The power of the imagination rather than the power of law or morality may be said to govern ideal theory. Ideal theory is an invitation to think of ourselves and our institutions in a specific way. In that quality, it can help us reinvent ourselves in the midst of our given societies: “... ideal theory, which defines a perfectly just basic structure, is a necessary complement of non-ideal theory without which the desire for change lacks an aim.”¹³

Deliberative democrats largely evade discussions about ideal and non-ideal theory by shunning positive substance in their largely procedural understanding of political justice, and hence the need to spell out an end state. They rather reconstruct which (often counterfactual) idealizations with regard to the possibility and desirability of mutual agreements we need to make if we want to respect ourselves as rational agents who co-govern their common world. Again, we see the idealistic move of the *as if* at work: *if* we look at ourselves and each other in this or that way, *then* certain obligations, rights, and expectations are implied. The strength of the *as if* is the strength of binding ourselves through imagining ourselves to be what, in real life, we cannot completely be: autonomous, free, equal agents, joined in forms of social cooperation for our mutual benefit. But

¹³ Rawls, *Political Liberalism*, p. 285.

we can build institutions that may help us approach conditions that enable us to act from reciprocal *as if* relationships.

§.4 – Bringing it back to the practice of political philosophy

At the beginning of this paper, I said that I wanted to bring together two things that are not usually brought together: the idea of a strongly embedded political philosophy and the attractions of ideal theory as a methodological way of bringing out radically new visions of the embedding society. I hope to have shown that these two things can be combined without contradiction. In the course of making these two points we have learnt that both a firm understanding of the rootedness of normative political theories and an understanding of ideal theory as an exercise in the philosophy of the ‘as if’ can help us prevent what Skinner called becoming ‘bewitched’ by our received ways of thinking.

First, acknowledging what I have called the rootedness of the theories that we work with makes us aware of their limitations. Paying attention to the kinds of questions that they were once meant to help answer may make us see that they are not ideally fit to be applied to our concrete questions in the here and now. We may have to open up our inherited philosophical framework in order to reach better philosophical judgment. Second, we have learnt that ideal theory may be one attractive way of going beyond the bewitchment of received ways of thinking of various perspectives in society and invite such perspectives to look at themselves from a new and possibly shared perspective.

What both moves have in common is that they enable us to see things in a fresh light and prompt us to keep thinking for ourselves rather than follow frameworks and expectations bequeathed on us through tradition, education, and what have you. I see no reason why in our political thought we would have to choose for either a Skinnerian openness to different intellectual traditions, or an ideal theoretical, or a non-ideal theoretical approach. I would rather say that much depends on the situation and on the question which kind of political theorist one is. In the practice of engagement in political theory, all these methodological approaches can have their value. And this not in the least because they can help disclose each other’s’ dogmatic blindness. Where the ideal theoretical mind will easily run the risk of seeing and understanding the world through the glasses of its own idealizations only, the non-ideal theoretical mind may lose sight of the grander vision of what we could achieve if only we would dare to think of ourselves under the aspect of some attractive philosophy of the *as if*. I have the strong suspicion that the best political theorists are capable of a certain flexibility when it comes to their choice

for these and other methodological approaches, and let it be determined by case-to-case needs rather than dogmatic belief.

2. The Importance of Ideal and Non-Ideal Theory

WORKING PAPER by Simon Rippon and Miklós Zala (Central European University)

Introduction

In recent years, mainstream academic political philosophy has come under increasing criticism for its relative neglect of broader contemporary political debates and real-world injustices (Anderson 2010; Farrelly 2007; Geuss 2008; Mills 2005; Sen 2006). While political philosophers have been quibbling with one another in more or less arcane journals about the nature of a utopian ‘just society’, political philosophy as a whole has been disengaged from both the questions and the kinds of views that animate policymakers, the media and voters, and is consequently lacking in political influence. If we accept that one important purpose of political philosophy is to use whatever insight can be gained from philosophical reflection to positively influence political discourse and political behaviour (beyond the walls of the seminar room), this gives us strong reason to refocus political philosophy to better address issues of justice that matter on a practical level.

We have no substantial disagreement with these views. We believe that (some) political philosophy ought to have a significant impact on the outside world. We accept the need for mainstream political philosophy to focus more thought on real-world injustices and on better realizing justice under the difficult and complex circumstances that we find ourselves in – what is known as *non-ideal theory* in the philosophical parlance. However, we do take issue with what we see as a tendency to throw out the baby with the bath water. Some philosophers have argued not just that there should be more work devoted to non-ideal theory, but that we have no practical need for its converse, *ideal theory*; the kind of thought about justice that abstracts or idealizes away from actual circumstances. What kind of principles of justice would be followed in a perfectly just society? What principles would govern a ‘well ordered’ society in which everyone accepts (and knows that everyone accepts) and complies with the principles of justice which govern basic social institutions? (Rawls 1999a, 5). What would be the principles of justice if we make the simplifying assumption that nobody is dependent on others for care, and all are equally capable of rationally pursuing autonomous, independent lives? (Rawls 1999a, 11). These are all questions of ‘ideal theory’, understood in a few of its various senses.

Some influential critics of ideal theory argue that reflecting on the ideal is irrelevant to the goal of identifying how to improve the level of justice in our society (Sen 2006, Wolff 2015). For example, according to Amartya Sen:

The question ‘What is a just society?’ is ... not a good starting point for a useful theory of justice ... it may not be a plausible end point either. A systematic theory of comparative justice does not need, nor does it necessarily yield, an answer to the question ‘What is a just society?’ (Sen 2006, 226).

What Sen means by ‘comparative justice’ is simply making judgments about better and worse. Thus, according to Sen, we can adjudicate more just, or less just, states of affairs without needing a theory of the *fully* just to do so. On Sen’s conception of ‘transcendental’ or ideal theory, it delivers something like a picture of an imaginary perfect island. But that perfect island is almost certainly unreachable, and a focus on it is simply not useful if our aim is to figure out how to improve the situation that we actually find ourselves in.

Others claim that ideal theory is even worse than useless. According to Elizabeth Anderson (2010, 5), the simplifying assumptions of ideal theory are *epistemically disabling*, blinding us to certain kinds of important injustice, such as race-based injustice. Charles Mills (2005) argues that ideal theory is *ideological* in the sense that it reflects and perpetuates the group privilege of the philosopher-as-bourgeois-white-male, enabling us to continue ignoring the voices of the oppressed.

We believe these criticisms deserve to be taken seriously. As we explain below, we think they provide a very strong case against pursuing ideal theory to the virtual exclusion of non-ideal theory, as has too often happened. But by defending the significance of ideal theorizing as an important *complement* to non-ideal theorizing, we seek to clarify its point and value. Contrary to Sen, we claim that a vague ideal of the just is at least implicitly appealed to in making comparative judgments about justice of the kind that Sen wants to make. And we think that thinking about the just society, thinking about something rather like the perfect island (but not quite like that, as we explain below), can help us to refine the vague ideal that we necessarily appeal to in making those comparative judgments.

We develop our argument as follows: In section two, in order to elucidate the concept of ideal theory more precisely, we examine the distinction between ideal and non-ideal theory. In section three, we rehearse and classify various criticisms of ideal theory that we have come across. In section four, we offer an argument explaining how ideal theory is useful for a non-ideal theorist, and identifying certain kinds of misuse of ideal theory. In sections five and six, we give two examples of how doing ideal theory can make a positive difference to our judgments in non-ideal theory. Finally, in section seven, we return to the criticisms of ideal theory and respond to them.

Our main point is of *practical* importance for the methodology of political philosophy and for its effectiveness in changing the world for the better, because if we are right, non-ideal theorizing would be greatly diminished by a failure to continue to attend also to ideal theory.

1. What is the distinction between ideal and non-ideal theory?

There is no single, clear distinction between ideal and non-ideal theory, as these terms are used in the literature (Valentini 2012; Hamlin and Stemplowska 2012). According to Laura Valentini's helpful classification (2012), we can understand the division between ideal and non-ideal theory along three distinct dimensions: *full vs. partial compliance*; *utopian vs. realistic*; and *end-state vs. transitional* theory.

On the first dimension, the 'full compliance' version of ideal theory designs principles of justice that assume that every, or almost every, individual complies with them (cf. Rawls 1999a, 8); its non-ideal 'partial compliance' counterpart aims to take into account the kind of unjust behaviour that we observe in the real world (to take an example, non-ideal principles might require individuals to do more than their fair share to realize justice in case they need to make up for others who do not do enough (cf. Miller 2011)).

On the second, 'utopian' vs. 'realistic', dimension, we can identify three distinct positions. A categorical 'utopian' ideal theory treats justice as a timeless ideal value, and is fact-insensitive, proceeding with theorizing about justice without regard to feasibility constraints (Cohen 2003). A 'realistic utopian' theory, in contrast, attempts to take into account facts about the human condition such as moderate scarcity and limited altruism, and is thus more closely linked to what we should *do* than to how we should think about value (Rawls 1999a, 126-130). But though a 'realistic utopian' theory is fact-sensitive in some respects, it assumes away at least some facts characterizing real-world politics, and to that extent it is an example of ideal theory. Those who reject it in favour of its 'realistic' non-ideal counterparts consider it necessary to take into account additional facts that the utopian theory idealizes away (Farrelly, 2007).

The third dimension of the ideal non-ideal distinction concerns whether the theory is an 'end-state' (or 'transcendental') theory, setting out a conception of the just society that could function as a long-term goal for institutional reform, or a 'transitional' (or 'comparative') theory, indicating how justice might be improved. On Rawls's reading, which presupposes end-state theory, a transitional theory is just a theory of how we should work towards the just end-state (Rawls 1999b, 89-90). Those who reject ideal end-state theory, however, claim that a transitional theory needs only to tell us how to make the world more just, and that this does not presuppose an understanding of what would make the world fully just (Sen 2006).

In this paper, we will understand ideal theory as a theory of justice that is ideal along one, two, or all three, of these three dimensions.

2. Against ideal theory

Let us now examine the arguments against ideal theory in more detail. We classify the criticisms of ideal theory under the following five headings:

a) It is infeasible

The charge of infeasibility begins from the observation that ideal theory simplifies or abstracts from the complexities of the real world. It may therefore make assumptions about facts that do not accurately reflect how the world is, or unrealistic assumptions about people's motivations. The result of these simplifications may well be that the 'just society' pictured by ideal theory may well not be a place we can get to from here. And its inaccessibility means that the principles of ideal theory are infeasible in practice (Farrelly, 2007).

b) It is dangerously impractical

The charge here is that if we take a utopian vision as our goal, when faced with the complex circumstances and difficulties of the real world, our attempts might well backfire. That is: aiming at a conception of a just society might actually make things worse. Suppose, for example, that our ideal theory includes very liberal freedoms of movement and of association, so that people can define and construct value in their lives in part by choosing to freely live in certain neighbourhoods and build valuable lives as embedded members of communities that share many common interests. Suppose, in the light of this view of what justice requires, we now accept these principles of free movement and association in the real world, perhaps even enshrining them in our constitution. What we end up with, in the real world, may well be persisting racially segregated neighbourhoods, and all their attendant injustices (Anderson 2010, 5-6). So, trying to apply a utopian ideal theory instead of focusing on what we should do here and now might well be an impractical approach.

c) It does not address our actual needs

A third criticism is that ideal theory does not give us useful, practical guidance that addresses our actual needs. Ideal theory focuses on describing a distant vision of value, a ‘perfect island’. But what we want to know is how to act in the circumstances we find ourselves in, and how to make things better, not just how to think about what perfection would look like (Sen 2006).

d) It diverts our attention, and is ideological

A different, stronger version of the third criticism is that, like a dream world into which we choose to escape, ideal theory actually diverts our attention away from many practical real-world problems that anyone concerned about justice should be focused on. According to Elizabeth Anderson, ideal theory blinds us to certain injustices. For example, the history of racial injustice is invisible to ideal theory, and is not grappled with, because on the ideal theoretical level, such injustice simply not exist (Anderson 2010, 5). The most stringent version of this critique says that ideal theory is ideological (Mills 2005). On this view, it is not a mere coincidence that ideal theory diverts our attention: rather, ideal theory is itself part of a power structure that reproduces injustice by privileging the concerns of the bourgeois white males who dominate philosophy as they dominate society in general (Mills 2005). Ideal theory is, on this view, little more than a rationalization for ignoring racial and gender injustice and silencing the voices and concerns of the oppressed (Mills 2005).

e) It implausibly appeals to timeless moral truths

Finally, ideal theory is sometimes criticised for some would describe as an implausible appeal to timeless moral truths (Geuss, 2008). Consider, by way of contrast, a view known as ‘political realism’. According to this view, assessing what ought to be done in politics is not a matter of engagement with normative truths that come from something beyond the internal norms and purposes of institutions and practices as they exist in each society. To assess what a government ought to do, or how it might govern better, we must rather examine the socially accepted practice and function of the government, and reflect on its internal standards. Some non-ideal theorists, attracted to this sort of view, think that ideal theory is implausible because it does not examine the political practices and institutions we actually have and analyse their internal norms, but rather seems to demand an appeal to timeless Platonic, moral truths fit to govern all possible political practices and institutions. And such truths seem doubtful, both on metaphysical and epistemological grounds. What are these eternal truths like if they exist ‘out there’, as if floating somewhere outside time and space. And even if they do exist, how on earth could we know about them?!

In contrast with ideal theory, advocates of non-ideal theory argue that it suffers none of the above weaknesses. It can offer feasible principles, since it is capable of taking into account the circumstances and limitations we are faced with in the real world. Second, in contrast to the utopian, sometimes dangerously impractical approach of ideal theory, it opens the way to a certain conservative pragmatism or incrementalism: we can aim to make things just a little better and more just than they are, rather than aiming at perfection and risking everything. Third, non-ideal theory addresses our actual needs and problems, paying attention to circumstances (and to injustices in particular) as they exist in reality. Finally, political realists believe that there is no need for non-ideal theory to make an implausible appeal to eternal moral truths. We can simply refer to the functions and internal standards of the social institutions and practices that we already have.

3. What use is ideal theory?

According to its critics then, ideal theory is infeasible, dangerously impractical, fails to address our actual needs, diverts our attention, functions as ideology, or appeals implausibly to timeless moral truths. If the critics are right, we should give up on ideal theory. On some of these views, ideal theory is not just useless, but positively harmful. We think these critics are wrong: we think ideal theory is useful, perhaps even essential, for serious thinking about justice.

As noted previously, Amartya Sen seems to endorse a version of the objection that ideal theory does not address our actual needs. As Sen puts the point, making judgements of ‘comparative justice’ does not require a ‘transcendental’ approach to justice, identifying perfectly just societal arrangements. That is, judging that one state of affairs is more just than another does not require us to know what ‘perfect justice’ consists in. By way of analogy, Sen writes:

We may indeed be willing to accept, with great certainty, that Everest is the tallest mountain in the world ... but that understanding is neither needed, nor particularly helpful, in comparing the heights of, say, Kanchenjunga and Mont Blanc. There would be something very deeply odd in a general belief that a comparison of any two alternatives cannot be sensibly made without a prior identification of a supreme alternative (Sen 2006, 222).

Sen is quite right, of course, that we don’t need to know about Everest to compare the heights of Kanchenjunga and Mont Blanc. And we should accept that the general principle Sen questions by using this example is false. It is obvious enough that comparisons of tallness do not require identification of anything maximally tall, comparisons of weight do not require identification of anything maximally heavy, and so on. But what conclusion should we

draw from this? From the fact that a general principle is false, it does not follow that none of its instances are true. It's false that all sheep are white, but it doesn't follow that *no* sheep are white. Our concern here is not about the general principle of whether comparative judgments require a transcendental approach, but about the specific case of justice: is it helpful to know what a perfectly just society would be like, if you want to adjudicate which is the more just of two states of affairs? More generally, is it helpful to develop an understanding of ideal theory, if you want to do non-ideal theory?

We think the answer to these latter questions is 'yes'. Our conception of what counts as an *improvement* in justice – what counts as a *more just* society – would be very limited, we believe, if we lack a grander view of what counts as justice *simpliciter*. And that pushes us towards a transcendental approach, a theory of the fully just society, or in other words, an ideal theory of justice. The widespread interpretation of Sen's point as a valid objection to ideal theory, we believe, stems from a misunderstanding of what ideal theory is.

What is ideal theory, then, in contrast to what Sen may have in mind? Ideal theory of justice, we believe, is not comparable to a photograph of a perfect island; it is not merely an image of the perfect, just society. It is, rather, a conception of a certain kind of excellence of social institutions.¹⁴ By way of analogy, let us consider some different ideals. Consider ideal health, for example. We can intuitively identify (more or less) what counts as ideal health. It is not unusual for doctors conducting routine check-ups to tell patients that they are in 'excellent health'. That does not mean we know what strictly *perfect* health is, or to have an *image* of perfect health. Ideal health might well be nothing more than the absence of diseases, and if it is, then it would be impossible to give a positive characterization of it. Thus, knowing what ideal health is not like knowing that Everest is the highest mountain. It is, rather, having a fairly vague conception of a certain kind of excellence. To take a second example, imagine the concept of the ideal meal. What would an ideal meal be like? It would be an unsurpassably excellent meal: it would have variety, lots of flavour, it would be nutritious, it would be just filling enough – it would exhibit a package of all the virtuous qualities that an excellent meal could exhibit together, with no shortfall in any of

¹⁴ Compare John Rawls' description of justice as "the first virtue of social institutions" (Rawls 1999, 3).

them. That's not to say that when we have a conception of the ideal meal we have one particular image in mind of a *perfect* meal, as if we could select one meal above all others as the single standard of excellence. These examples suggest that Sen's mountains might be a poor analogy for ideal and non-ideal theories of justice. Having a conception of the ideal of justice is not like knowing *that Everest is the highest mountain* but is more like knowing what *the high* is, or *understanding* what the high is (the analogy runs out here because height seems like too simple a concept to demand very much in the way of understanding).

Now, to make the discussion of ideal theory of justice more concrete, let's take a specific example: John Rawls's theory of *justice as fairness* (Rawls 1999a). When critics have attacked ideal theory, Rawls has often been their target. But Rawls does not give us anything analogous to a mere picture of a perfect island. Rawls's theory does, indeed, give us a set of principles for a just society: the principle of equal basic liberties, the principle of fair equality of opportunity, and the difference principle. But Rawls's theory is not just these principles of justice; it is far richer than that. In addition to the principles of justice, Rawls's theory marshals a set of reasons and arguments. These justify and explain the significance of the principles of justice, and their lexical ordering. They provide us with a particular method for arriving at principles of justice, and a justification for using it: the method of reflective equilibrium, and the thought experiment of choosing principles for the future organizing society from the original position, behind the veil of ignorance. Thus, what Rawls's ideal theory of justice gives us is not just an image of transcendental 'perfect justice', but a far richer, more developed understanding of excellence of social institutions. Of course, Rawls has many opponents who disagree substantively with his conception of justice. But a substantive disagreement about the principles of justice, or how they are arrived at, is not an argument against ideal theory.

In summary, Sen seems right about the general principle that making comparative judgments does not necessarily require a transcendental approach. But this can only be interpreted as an objection to ideal theorizing about justice if we have a very anaemic and unrealistic conception of ideal theory in mind: a conception according to which ideal theory provides us with no more than an unexplained image of perfection.

Consider the following argument, made by Sen:

Identification of the best does not specify a unique ranking with respect to which the best stands at the pinnacle; indeed the same best may go with a great many different rankings with the same pinnacle. To consider an analogy, the fact that a person regards the Mona Lisa as the best picture in the world, does not reveal how she would rank a Gauguin against a Van Gogh. (Sen 2006, 221)

Sen is quite right to say that a simple identification of the Mona Lisa as the best is not going to determine how to rank a Gauguin against a Van Gogh. But neither, for that matter, is it going to determine how to rank a Gauguin against the scribbles of a kindergartner. The problem here is not just that a conception of the best does not necessarily determine a theory of the second best. The problem is that as Sen conceives it, transcendental theory appears to be a very anaemic conception of excellence. To take the analogy a bit further, suppose we are in very ‘non-ideal’ artistic circumstances, producing kindergarten-level stick figures, having never seen, or developed the necessary skills to produce, anything better. Suppose we are now shown the Mona Lisa and told ‘This is better. In fact, this is the very best. Do your art like that!’ This is not likely to be useful advice, since it is way beyond the limits of feasibility. Moreover, merely learning that the Mona Lisa *is* the best does not give us any understanding of *why* it is best. Providing such an understanding, we believe, is an essential part of the role of ideal theory. Suppose that instead of being shown the Mona Lisa, we become autodidacts, making merely comparative judgments about our present output and the nearby possibilities. Perhaps we find that our stick figures are improved if we clearly depicted expressions of happiness or sadness on their faces. Now we have found a practically useful conception of betterness for our current circumstances. But it will be immediately obvious to the astute reader how limited, and how potentially misleading, this conception of betterness is likely to be. The rule that distinguishes the better stick figures, ‘It is better to clearly depict the emotion on the face of the figure’, is totally misleading when extended beyond our current circumstances: it is incapable of giving us, for example, any grasp of the excellence that those who understand something about artistic virtue recognize in the intriguing ambiguity of the Mona Lisa’s ‘smile’. The imagined rule that underlies our comparative judgment – show the emotions on the figures’ faces – is in fact extremely limited in application. But unless we develop a general understanding of the *excellences* of art, it is difficult to see how we could possibly recognise the limits of application of this rule. With nothing to guide our very limited comparative judgments beyond the judgments themselves, it is difficult to see how we should continue to make reliable comparative judgments. And this makes it difficult to see how we could make any significant progress without error. By providing a general understanding of excellence, ideal theory may enable us to make greater progress without falling into error.

The foregoing argument justifies interest in a ‘transcendental’ theory of justice. But in section 2, following Valentini, we identified three dimensions in which a theory of justice may be more ideal or less ideal. The first dimension was end-state vs. transitional, or ‘transcendental’ vs. ‘comparative’ approaches in Sen’s terminology. The other two dimensions were full vs. partial compliance, and utopian vs. realistic fact sensitivity. Both these dimensions of ideal theory concern not just a focus on an ultimate normative ideal, but bracketing or simplifying assumptions about real world facts. Can we justify idealizing also in these respects, if our interest is in developing a theory that will inform us about the excellences of justice? We believe so.

To argue for this claim, we borrow an analogy between ideal scientific theory and ideal theory of justice drawn by philosopher of science Jenann Ismael (2016, 14-18). Ismael compares Rawls's theory of justice with Newton's model of a 'simple' or 'ideal' pendulum. The ideal pendulum is a weight that is not subject to air resistance swinging from a frictionless pivot by a massless cord. These unrealistic assumptions mean that no real-world pendulum is exactly like an ideal pendulum. But Newton saw that the behaviour of an ideal pendulum could be modelled by a simple equation – the Law of the Pendulum – which directly relates the pendulum's period to the length of the cord, for a given strength of gravitational field. Thus, Newton's idealization helps isolate and clearly expose the effect of the force of *gravity* on a pendulum, helping Newton derive his theory of planetary motion, and helping everyone better understand what gravity does, and thus what it is. Idealization in physics helps us set aside exogenous influences that greatly complicate the behaviour of real world objects, in order that we can more clearly identify fundamental forces. Similarly, the simplifications of ideal theory in political philosophy can help us more clearly identify the fundamentals of justice. As Ismael writes:

When Rawls wrote *A Theory of Justice*, and focused on the ideal part of the theory, he did not set out to produce what purported to be a model of the actual world and *fail* any more than when Newton described the behavior of an ideal pendulum, he set out to describe the behavior of an actual pendulum and *failed*. ... [Rawls's purpose] was to exhibit what a just society looked like in its purest expression, unobscured by noncompliance. It was also to show how the principles that organize such a society follow from the account of justice as rational choices made in an original position of equality. The full development of a complete theory of justice should have the resources to deal with the effects of noncompliance, but the ideal part of the theory is the part of the theory that most clearly displays the *content* of, and *justification* for, his notion of justice. (Ismael 2016, 16).

Ismael's point is that ideal theory, in justice as in science, enables us to set aside some of the complex interactions that take place in the real world, in order to identify and model the fundamental forces, norms, or relationships we want to think about all the more clearly.

Building the analogy further can also help us understand why some of the objections to ideal theory are misplaced, or rather should be couched as objections to a certain kind of *misuse* of ideal theory. Since ideal theory does not attempt to accurately reflect the real world, it is no objection to it that it does not do so, and is thus infeasible, or that it 'ignores' certain kinds of injustice. Nor is it an objection that someone who applies ideal theory without thought for the differences between the ideal model and the real world is likely to end up with bad results. Suppose that, aware of Newton's theory of the ideal pendulum, we now wish to build a machine involving a real pendulum with a specified period, such as a grandfather clock: this is an engineering problem. If

we try to solve it simply by applying the ideal theory, we will fail. Our real pendulum will be affected by friction and air resistance, and thus will eventually stop unless energy is imparted to it. Its cord will inevitably have mass, so its period will be affected by the relationship between the mass of the cord and of the weight. The cord will lengthen or shorten according to temperature, changing the period. In short, if we attempt to unthinkingly apply our ideal theory to our engineering problem, ignoring the complications of the real world, we will certainly not do well. This does not imply, of course, that ideal theory in physics is of no practical value when it comes to engineering problems. We can perhaps imagine people who work entirely ‘bottom up’ as engineers guided by intuition and by trial-and-error, building machines without any grasp of the mathematics that describes the underlying physical principles. But of course, being able to mathematically understand what is going on in our ideal physical model helps us understand better what is going on in the real world, and so helps us build better and more sophisticated machines than we would otherwise have been able to build. Our ideal physical models cannot be applied straightforwardly to the real world, and so they do not obviate the need for getting our hands dirty doing engineering. But they are nonetheless useful. Analogous claims hold for the ideal theory of justice, which stands to improving justice in the real world as ideal theory in physics stands to engineering. We should be wary of attempts to directly apply the prescriptions and principles of ideal theory in the real world. But an understanding of ideal theory can guide an intelligent person who seeks to bring about improvements in justice in the real world, just as an understanding of physics can guide an engineer.

As Ismael rightly points out (2016, 28), the point here is not that we need a complete and final ideal theory of justice before we can do non-ideal theory, any more than we need a final theory of physics before we can start doing engineering. We can pursue both in simultaneously, with the one informing the other.

4. Border Justice

If the analogy drawn with physics and engineering and the other conclusions of the previous section are correct, then the pursuit of ideal theorizing about justice should inform and influence our non-ideal theorizing; illuminating and clarifying the normative standards we posit. Even political philosophers who are critics of ideal theory have spent a great deal of their professional lives thinking about it. We think that some observations about how the judgments of political philosophers differ from those of the general public, and the kinds of reasons they would give in support of those differing judgments, shows the influence of ideal theory on their thinking.

Consider the non-ideal theoretical question of how Europe should deal with the wave of refugees and economic migrants flowing across its borders from Africa and Asia in recent years. Now notice that the active

debate about justice among the general public, the media and political leaders is almost entirely constricted to an argument about the rights of refugees. There is almost no argument about the claims of economic migrants, since even after separating out legal issues, it is almost universally assumed that this group has no valid claim of justice against European citizens and governments. A commonly accepted view among the general public is that these foreigners are coming over to take what is ours, taking advantage of generous European social benefits and high wages. Europe should build up its hard border defences, and while it's at it, stop wasting so much taxpayers' money on foreign aid.

What is striking is how rare such a view is by contrast among political philosophers, even among those who are comparatively politically right-leaning, and those who argue in defence of border controls and distinctions between the legal rights of citizens and non-citizens as just (Blake 2003; Miller 2005; 2016). Most political philosophers consider that national borders make arbitrary distinctions between the rights of persons, and that citizenship is by and large unearned. Philosophers tend to advocate tear down the walls, or opening immigration systems, or at the very minimum, massively increasing foreign aid so that European citizens and foreigners enjoy separate, but far more equal, opportunities than are currently available.

It is easy to overlook the fact that *both* views can stem from theories of justice. Let us examine first the view that is widely accepted among the general public. There is a reasonable position that generates this view: It is unjust for people to help themselves what they are not entitled to, and foreigners are not entitled to what is ours. As the saying goes, 'charity begins at home', and our duties of justice and of beneficence are to our fellow citizens first and foremost. So, justice requires that we protect and defend the benefits of our societies for our fellow citizens.

The view widely shared among political philosophers, on the other hand, stems from a different theory of justice; one that we believe is far more general and grounded in more ideal thinking about justice. That is, we reach it by thinking a bit about what an ideally just society would look like. When most political philosophers look at the plight of economic migrants pushed back at national borders, what they observe is coercion (of some but not others), a stark difference in distributional entitlements recognized by governments, and stark inequalities between European citizens and most would-be economic migrants from Africa and Asia. They then connect this observation with a general claim that *arbitrary inequalities are unjust*. To assess the claim that citizenship justifies these inequalities, they consider whether citizenship is itself a morally arbitrary distinction between persons. And finding it morally arbitrary, because it is unchosen, or a matter of brute luck, or whatever, most philosophers conclude that the *status quo* leaves these would-be economic migrants subject to great injustices. The point we want to make is that philosophers do not simply question the entitlement that is connected to citizenship on its

own and intuit its injustice. The typical philosopher's view stems from a much more general view about *what a just society would look like*. In particular, a just society, most philosophers believe, would be a society in which morally arbitrary inequalities such as where one happens to be born do not make a huge difference to one's life prospects. The typical philosophers' views about the European border crisis are thus informed and influenced by some very abstract, ideal theory.

5. Racial segregation

We now offer a second example of how ideal theory can be instrumental in thinking about real world injustices. We here examine two approaches from the recent literature in political philosophy on the problem of racial segregation: those of Charles Mills's (2017) and Tommie Shelby (2016).

Charles Mills is a prominent critic of ideal theory. While he attacks ideal theorizing from several angles, his main criticisms are that ideal theory is inherently ideological, perpetuating white, male privilege, and that it is useless for corrective justice because its starting point -- a society devoid of racial injustice -- is unrealistic (for the claim that ideal theory is white ideology, see Mills 2005; for the claim that ideal theory is fundamentally flawed because it cannot come up with principles of corrective justice, see Mills 2013). Mills himself proposes a positive theory of racial justice which he regards as free of the above-mentioned shortcomings of ideal theory (Mills 2017, 201-16). Mills's positive theory, however, is surprisingly permissive to ideal theory, or to put it differently, it is surprisingly ideal-theory-like. Mills basically modifies Rawls's three principles of justice (the equal basic liberties principle, the fair equality of opportunity principle and the difference principle) in order to correct 'illicit white advantage/white privilege' (Mills 2017, 211). The result is that Mills keeps Rawls's first principle of justice, but the basic liberty principle should now take illicit white advantage into account. Mills accepts only formal equality of opportunity (again, with the distinction that it should aim to correct for white privileges), and he replaces the difference principle with the principle of 'Respect', which 'is included as a basic social good in keeping with both Kantian and Rawlsian norms, and the need for correcting the founding of the polity on the systematic disrespect, dissin', of people of color' (Mills 2017, 211). Mills also drops Rawls's lexical ordering of the principles of justice, admitting that he is uncertain about the ordering of his principles (Mills 2017, 211). In addition, Mills accepts what he describes as 'the essence of Rawls'. This is the hypothetical choice situation from the original position behind the veil of ignorance that Mills holds, despite all his criticism towards Rawls, to be 'a significant contribution to political philosophy' (Mills 2017, 212). Thus, Mills uses the same Rawlsian contractual apparatus with a twist, because the task is not identifying principles of distributive justice for an ideal society, but 'seeking principles of corrective justice for a non-ideal (here racist) society' (Mills 2017, 213).

Mills's positive theory, for an outspoken critic of ideal theory, is thus very surprising, because he accepts what is perhaps the most abstract element in Rawls's theory (the hypothetical choice situation), and he also accepts that something like the equal liberty principle and equality of opportunity must be principles of justice. The real difference is the need for correcting historic injustice, which was deliberately not Rawls's aim – Rawls set that aside for a later development of non-ideal theory. But, for all his criticism of ideal theory, it is clear that Mills's own proposed theory is just a modified social contract theory that aims to correct the injustices of the past and their remnants. Mills thus grounds his non-ideal political philosophy in a theory that is ideal in at least two of the dimensions described in section two: it is an end-state theory that also makes unrealistic assumptions, and thus is utopian.

In contrast to Mills, Tommie Shelby fully accepts that the Rawlsian ideal theoretical apparatus can be useful for theorizing the problem of social segregation and believes that applying Rawls to racial segregation is not a misconceived project (see Shelby 2013; 2016). Shelby thinks that Rawls's idealization is not a 'whitewashing' of history, but a move that is necessary to generate impartial principles that can be generalized to other cases (Shelby 2013, 150). In Shelby's view, the fact that the participants of a Rawlsian social contract do not know the colour of their skin is 'a way of modeling the *impartial* point of view of political morality, a way of preventing particular social contingencies—including *existing and past injustices*—from biasing the chosen principles or undermining their generality' (Shelby 2013, 150).

Shelby regards ideal theory as incomplete, but as still providing standards against which we can measure our real-world circumstances. He claims that Rawlsian fair equality of opportunity is useful for mitigating race-based injustice and disadvantage, because this requires forward-looking, and not only backward-looking principles, and the Rawlsian framework is able to provide those (Shelby 2013, 158). He also draws far-reaching conclusions from the fact that due to the failure of the state, Rawls's basic assumption of reciprocity is not satisfied (Shelby 2016, 212-9). Thus, Shelby uses Rawlsian justice as fairness as a standard to which the current, unjust state of affairs can be compared. Shelby (2016, 19-22), for example, claims that Rawls's emphasis on the importance of the basic structure motivated his view that racial segregation is a systematic injustice against blacks in the US. Moreover, as we mentioned above, social segregation definitely violates Rawls's fair equality of opportunity principle, and arguably it violates the Rawlsian ideal of distribution for the social bases of self-respect as well. In addition, Shelby uses Rawls's idea of reciprocity to point out that the lack of reciprocity that characterizes ghettos is unjust and it has consequences what kind of allegiance ghetto residents are owed to other citizens (Shelby 2016, 219-223).

We hold that Shelby's analysis provides a good example of how ideal theory can be very useful for analysing real-world injustices and guiding social reform, even while it is not to be understood as a concrete analysis or

diagnosis of actual injustices, or as a blueprint for public policy that one should unthinkingly implement. There might be a considerable gap between the principles of ideal theory and the task of finding solutions to real world injustices. Thus, ideal theory certainly needs supplementing with non-ideal theory. But it simply does not follow that ideal theory is of no use. And these examples, we hope, show many of the ways in which ideal theory is practically useful when it comes to thinking about and responding to real-world injustices.

6. Responding to the objections

Having explained how we conceive of ideal theory and its use in identifying and working toward solutions for real-world injustices, we return to the criticisms of ideal theory enumerated in section three. Recall that it was claimed that ideal theory is infeasible, dangerously impractical, fails to address our actual needs, diverts our attention, functions as ideology, or appeals implausibly to timeless moral truths. Having explained and given examples of what ideal theory is and what, in our view, it is for, we now have the materials in place to respond to all but the last of these criticisms.

First, regarding the criticism that ideal theory is infeasible, we would like to emphasize that just as engineering does not stop at physics alone, deciding what to do in the real world cannot stop at ideal theory alone. Like an ideal model in physics, ideal theory is not intended to be applied directly to the complex reality, but to inform a deeper understanding of fundamental norms and relationships which guides the making of practical judgments about what to do.

There are a couple of slightly different versions of the infeasibility objection that we did not mention earlier. One of them is that since there is deep disagreement about what ideal theory we should accept, we cannot possibly reach a final version of ideal theory that we can use to solve practical problems. But the foregoing discussions have shown why we do not need a final ideal theory, any more than we need a final theory of physics to do engineering. Moreover, this criticism is vulnerable to a *tu quoque* objection: there is as much disagreement about non-ideal theory as there is about ideal theory.

The third version of the infeasibility objection is that citizens under ideal theory would have to have unrealistic motivations, whereas we need in David Miller's memorable phrase, 'justice for earthlings' (Miller 2013). In response, we hold that ideal theory can be very demanding, but even an ideal theory should be realistic *enough* to still be an account of justice for human beings. It must be something human beings *can* live up to. For example, Rawls's ideal theory of justice as fairness requires virtuous motivation, but (arguably) non-heroic virtuous motivation (Rawls 1999a, 112).

Regarding the criticism that ideal theory is dangerously impractical, we would like to refer again to the point above, pointing out that just as an engineer who attempts to build a purely 'ideal' machine based on models in physics will fail, so will political philosophy fail if it neglects attention to differences between the ideal model and the real world. There is an analogy here between John Stuart Mill's point, when he makes a distinction between treating the utilitarian principle as a procedure for deciding what to do (a *decision procedure*), and treating it as a theoretical determinant whether a given action is right or wrong (a *criterion of rightness*).[cite] As Mill points out, anyone who attempts to think about and follow the utilitarian principle of maximising the sum of human happiness in their everyday actions is likely to do very badly at maximising the sum of human happiness (they would be too busy calculating to do very much at all, for one thing!) Similarly, ideal theory cannot be simply a direct guide as to what to *do* here and now, but should give us a better understanding of what counts as just, and what count as improvements in justice.

Regarding the criticism that ideal theory fails to address our actual needs, we of course concede that it is incomplete, and that attention needs to be paid to non-ideal theory as well. The trend in political philosophy to concentrate almost exclusively on ideal theory is, we agree, a regrettable one. Similarly, for the charge that ideal theory is attention-diverting and ideological. Nothing that has been said recommends doing ideal theory on its own. There can be a division of labour between ideal and non-ideal theorizing and as we pointed out above, they can be and must be done simultaneously.

There is one objection that we have no dealt with yet; it is that ideal theory appeals implausibly to timeless moral truths. One of us shares doubts about the kind of Platonist metaphysics that lies behind this objection, and believes that timeless moral truths do not need to rest on such a metaphysics. But whether or not this view is ultimately correct, the political realist here faces a dilemma. On the first horn of the dilemma, the political realist restricts his criticisms of social institutions and practices to their accepted internal standards. On this horn, it seems that we can criticize the Nazi administration charged with deporting the Jews to concentration camps was defective only to the extent that it did not perform its function well. This horn of the dilemma involves an illegitimate leap from 'is' to 'ought'; from the claim that some standard is socially accepted to the claim that that standard ought to be obeyed. On the second horn of the dilemma, the political realist frees himself to appeal to additional standards external to the socially accepted practice itself. But on this horn of the dilemma, he is vulnerable to a *tu quoque* objection. For his appeal to external standards no less vulnerable to the anti-Platonist objections than those of the ideal theorist.

References

- Anderson, Elizabeth. 2010. *The Imperative of Integration*. Princeton, NJ: Princeton University Press.
- Blake, Michael. 2003. 'Immigration'. In *A Companion to Applied Ethics*, R. G. Frey and Christopher Heath Wellman (eds.). Oxford: Blackwell, 224–37.
- Cohen, G. A. 2003. 'Facts and Principles'. *Philosophy & Public Affairs* 31 (3): 211–45.
- Farrelly, Colin. 2007. 'Justice in Ideal Theory: A Refutation'. *Political Studies* 55: 844–864.
- Geuss, Raymond. 2008. *Philosophy and Real Politics*. Princeton: Princeton University Press.
- Hamlin, Alan and Zofia Stemplowska. 2012. 'Theory, Ideal Theory and the Theory of Ideals'. *Political Studies* 10 (1): 48–62.
- Ismael, Jenann. 2016. 'A Philosopher of Science Looks at Idealization in Political Theory'. *Social Philosophy and Policy* 33 (1-2): 11–31.
- Miller, David. 2005. 'Immigration: The Case for Limits'. In *Contemporary Debates in Applied Ethics*, Andrew I. Cohen and Christopher Heath Wellman (eds.). Malden, MA; Oxford: Blackwell, 193–206.
- Miller, David. 2011. 'Taking up the Slack? Responsibility and Justice in Situations of Partial Compliance'. In *Responsibility and Distributive Justice*, Zofia Stemplowska and Carl Knight (eds.). Oxford: Oxford University Press, 230–45.
- Miller, David. 2013. *Justice for Earthlings: Essays in Political Philosophy*. Cambridge: Cambridge University Press.
- Miller, David. 2016. *Strangers in Our Midst: The Political Philosophy of Immigration*. Cambridge, MA: Harvard University Press.
- Mills, Charles W. 2005. '“Ideal Theory” as Ideology'. *Hypatia* 20 (3): 165–83.
- Mills, Charles W. 2013. 'Retrieving Rawls for Racial Justice? A Critique of Tommie Shelby'. *Critical Philosophy of Race* 1 (1): 1–27.
- Mills, Charles W. 2017. *Black Rights/White Wrongs: The Critique of Racial Liberalism*. Transgressing Boundaries. Studies in Black Politics and Black Communities. New York, NY: Oxford University Press.
- Rawls, John. 1999a. *A Theory of Justice. Revised Edition*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Rawls, John. 1999b. *The Law of Peoples*. Cambridge, MA: Harvard University Press.

- Sen, Amartya K. 2006. 'What Do We Want from a Theory of Justice?' *The Journal of Philosophy* 103 (5): 215-38.
- Shelby, Tommie. 2013. 'Racial Realities and Corrective Justice: A Reply to Charles Mills'. *Critical Philosophy of Race* 1 (2): 145-62.
- Shelby, Tommie. 2016. *Dark Ghettoes: Injustice, Dissent and Reform*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Valentini, Laura. 2012. 'Ideal vs. Non-Ideal Theory: A Conceptual Map'. *Philosophy Compass* 7 (9): 654-64.
- Wolff, Jonathan. 2015. 'Social Equality and Social Inequality'. In *Social Inequality: On What It Means to Be Equals*, Carina Fourie, Fabian Schuppert and Ivo Wallimann-Helmer (eds.), 209-25.

3. Feminist Political Theory and Non-Ideal Theory as its Methodology

WORKING PAPER by Oana Crusmac (University of Bucharest)

In a previous paper I have looked into detail at the various implications the methodological debate on ideal and non-ideal theory launches. The main focus of the previous paper was to offer guidance on understanding what non-ideal theory is given that it does not have a clear definition and given that there is little consensus among political theorists on establishing what non-ideal theory exactly is. Unlike ideal theory, which was coined and sketched in detail by John Rawls, non-ideal theory was only slightly drafted by Rawls as he preferred to focus on the former and only mentioned some aspects of the latter. Rawls saw ideal and non-ideal theory as complementary, but the methodological distinction had also a ranking in his view: non-ideal theory had to rely on and to apply the principles identified through the lenses of ideal theory. As such, ideal theory gained, in Rawls' and his followers' perspective, a higher importance. Non-ideal theory had to be thought of only after principles of justice were issued in the framework of ideal theory. Disagreement on the latter (e.g. various types of distributive justice theories drafted in 'ideal' language) led to an even higher confusion on what is, what does non-ideal theory (since the ideal was hard to determine). Non-ideal theory has been, as I have pointed, erroneously equated with (1.) theory of second best alternative (derived from Sen's comparative account on justice theories); (2.) political realism; or (3.) empirical research as "action design and implementation" Robeyns (2008).

Moreover, as the term (non-ideal theory) was coined also by Rawls and put in a dependency relation to his account of ideal theory, the critiques addressed to ideal theory also correlated with non-ideal theory and this led to the main question of the previous paper: what is non-ideal theory and whether we should strictly understand it in the manner Rawls coined it. In the previous paper, I have showed that there are five possible interpretations on what non-ideal theory is, which can be grouped into two main categories:

A. Transitional non-ideal theory or rawlsian non-ideal theory: this interpretation derives non-ideal theory from ideal theory

(a.1.) Transitional non-ideal theory that follows solely short-term goals directly derived from ideal theory (see Simmons 2010);

(a.2.) Transitional non-ideal theory that follows only long-term goals that ends up as a completely unachievable utopia;

(a.3.) Transition non-ideal theory which tries to achieve the long-term goal through appropriate short-term goals (see Gilabert 2012).

B. Standalone non-ideal theory: this interpretation does not grant ideal theory a superiority or priority over non-ideal theory. As such, non-ideal theory is seen as detached, autonomous from the 'tyranny of the ideal'.

(b.1.) Comparative non-ideal theory that does not use any ideal or idealization: this interpretation corresponds to Sen's notion of comparative theory, where the second best option is chosen. The lack of clarity of this interpretation allows us to equate it also with implementation design or political realism.

(b.2.) Independent non-ideal theory that is independent from ideal theory but it constructs its own ideals based on a fact-sensitive methodology that allows developing principles by employing abstraction without idealization.

I concluded that only subtypes a.3. (which I consider the proper rawlsian interpretation of non-ideal theory) and b.2. represent viable accounts on how to understand non-ideal theory. Hence, we are left with two possible definitions of non-ideal theory. The transitional interpretation follows Rawls' definition of the complementary couple ideal - non-ideal, whereas the independent interpretation that uses abstraction to build its own principles is largely endorsed by critical race theorist and many feminists who claim that the fact-insensitive abstraction makes ideal theory either useless, either dangerous.

In this paper I will focus on the type of independent non-ideal theory supported by feminist political theorists such as Elizabeth Anderson (2009; 2010), Alison Jaggar (2009; 2015), Iris Marion Young, Lisa Schwartzman (2006a; 2006b). The approach shared by aforementioned authors relies on doing non-ideal theory by avoiding relativism through the method of abstraction.

The paper is structured as follows: in the first section I will focus on the critiques addressed to ideal theory that led to the birth of non-ideal theory conceived as completely detached and independent from ideal theory. In the second section I will present several feminist proposal of doing non-ideal theory, by relying on

abstraction without idealization and group relations methodology. The last section launches the question of whether we can equate independent non-ideal theory with critical theory.

Non-ideal theory as independent from ideal theory: the interpretation supported by theorists of the marginalized – feminists and critical race theorists.

This set of critiques does not follow the already classical refutation of ideal theory as unfeasible or impossible to guide action. As I have presented in detail in the previous paper related to the topic of ideal and non-ideal theory, rejecting ideal theory on the basis on the full compliance assumption is misleading, because the unfeasibility claim does not hold. As Estlund (2014) also pointed out, ideal theory is rejected mainly (the mainstream critique) because its ideals are impossible to achieve. However, there are two distinct interpretations of the meaning of ‘impossible’: ability (“is impossible for her to do so”) or probability (“it is impossible that she will do it”) (Estlund 2014, 119). Estlund considers, and I support his view, that impossible is different from zero probability, and that most of the principles issued by ideal theories are not impossible to be followed, but rather they have a low probability of being achieved. But, as he stressed, zero probability does not mean inability, and could only refer to a lack of will (Estlund 2014, 119). As such, the unfeasibility critique of ideal theory can be dismissed if we take into account his proposal - that of not looking for action-guidance in the principles of justice - as there is no mandatory requirement for justice to have immediate, practical value (Estlund 2014, 130-134).

Instead, the set of critiques I am focused on in this paper points out the issue of the idealized (on a white male pattern) moral agents. The autonomous, independent, rational individuals come in relation not only with the idealized methodology, but with the topics approached by ideal theory. As such, critical race theorist and many feminists have pointed out that a change in the content of theories of justice is needed: justice can not be approached solely from a distributive standpoint - where there can be only two elements of analysis: the metric/currency (‘what’) and the pattern of distribution (‘how’) - and proposed to also investigate aspects that are generally framed under the term ‘relational justice’, like oppression, domination, family, caregiving and dependency, etc.

As well as a change in the method through which theories of justice are coined is also needed. Even if these critiques of ideal theory reject relativism, they also reject that the principles of how a just society would look like to be invented in a solitary, monological and limited manner. Here also enters the critique of the autocracy of the political theorist drafting the theory. Instead, critical race theorists and many feminists propose to replace the monological method of ideal theory with a deliberative, non-ideal method. In other words they

propose to replace starting with the general in order to get to the particular with the opposite, namely to start from particular instances of injustice in order to achieve general justice. As Mills (2005) stressed, general principles issued in the framework of ideal theory abstract away from gender and race.

What has drawn my attention on the topic of non-ideal theory is fact that the critique addressed to Rawls' idealized agents of this Theory of Justice is shared by both critical race theorist and feminist theorists as well as by relational justice theorists. One can notice that theorists that represent "historically subordinated groups have always been deeply skeptical of ideal theory [...] and are attracted to nonideal theory, or what significantly overlaps it, 'naturalized' theory" (Mills 2005, 170). One possible explanation of this tendency is that "philosophers tend not to come from oppressed groups" (Cudd 2006 in Mills 2014, 34). Neither critical race theorists, nor feminists endorse a relativist account of justice, but they hope to construct a methodology that would help reaching objective principles of justice. The problem identified by these theorists that propose to replace ideal theory with non-ideal theory is that objectivity is not yet met in the classical literature: relativism is not objective, but neither ideal theory. For Mills (2005, 168), one of the main issues of ideal theory is that it relies "on idealization, to the exclusion, or at least marginalization, of the actual". This is why Mills (2005, 166) famously said that ideal theory is ideological and perpetuates group privilege.

The theorists that make the subject of this paper, who have dismissed the relevance of ideal theory highlight the main pitfall of ideal theory: that of the white male idealization. As such, some feminist political theorists (e.g. O'Neill 1987, 2000; Anderson 2010) consider that a proper form of abstraction represents the solution to both relativism (which would justify injustice) as well as to the 'tyranny of the ideal'. For many feminists and critical race theorists experience is essential when doing political philosophy: taking experience of (marginalized, ignored) social groups into account is seen as the solution to omission – for example instead of adopting color-blind ideals we should to add the experience of color, because color-blind policies reinforce the white bias. The same applies to gender inequality where if we ignore the lived experience of women we ignore also issues related to care and dependency.

Unfortunately, the idealized while male perspective still dominates in political theory: distributive justice theories focused on various patterns ('how') and currencies ('what' to distribute) still dominate the academia, white relational accounts, or care accounts are ignored or mentioned only as a secondary critique.

Feminists differ from the mainstream political theory in two aspects: the topic and the method. Both shifts proposed by feminists (topics and method) start from the 'who' of justice. In this paper the main focus will be the method, but one can easily notice that the shift in the method is strongly linked to the limitations of the topics covered by the mainstream political theory. When one theorizes on the distribution of goods/resources/etc

(quantifiable currencies) in order to have a just society, is quite easy to issue models while using some idealized assumptions of the person (as rational, self interested, independent, autonomous, etc). All the mainstream liberal theories of justice don't look at all at the relationship between individuals: these theories do not explore the relationship between the members of the family, they also ignore children, ill, disabled or mentally disabled people - as they cannot be account as independent and rational, nor does the caregiver count as self interested. This omission, non-ideal theorists argue, is given by blindness to particular lived experiences of some groups. Many political theorists that focus on marginalized groups, like people of color, women, dependants (children, ill, mentally disabled people), end up embracing non-ideal theory as a method that would allow including their groups in the topic of justice because ideal theory does not address their particular, group-based injustices. For example caregiving does not fit abstract liberalism's ideal theory because good caregiving is "other-directed", thus cannot fit the "self-interested" liberal action (Kittay 2001, 562). However, (and here lies the inquiry I will address toward the end of the article) the non-ideal theory explored in this paper does not contradict liberalism. While some feminist theorists (e.g. Schwartzman 2006a) that embrace non-ideal theory reject liberalism, others (Anderson 2010) consider that we can have a non-ideal theory conceived in a liberal framework.

In the next section I will present the methodological proposal of doing non-ideal theory that is sensitive to oppression. Elizabeth Anderson (2010), who embraces liberalism, proposed a route similar to the one advocated by MacKinnon (1989), who rejected liberalism on the basis that liberal objectivity fails to address substantive gender inequality. This form of non-ideal theory stipulates that the "theory must be built from bottom up, not from top down (Watson 2013, 35). Starting from bottom up means starting the lived experience of women and other oppressed groups excluded from the classical ideal theory. For MacKinnon and other feminists, "knowing the forms of injustice is crucial to remedying such injustice" (Watson 2013, 43).

As Srinivasan (2018, 596) points out, moral philosophy considers the female experience as belonging to the field of non-ideal, and since "feminism is deeply concerned with women's experience" this means that feminism is closer to non-ideal theory both in what regards its content as well as its method. The shift of feminist political theory from ideal theory to non-ideal theory can be explained also using Zerili's (2008, 106-107) four critical projects of feminism: (1) feminism tries to expose the absence of women from canonical discussions of politics; (2) feminist tries to integrate women in the categories of political membership from which they were originally excluded, (3) feminism shows that women cannot be integrated by adding and stirring as their exclusion was constitutive for those political categories, (4) as a result of the impossible inclusion of women, to reconstitute the categories of politics anew. Looking at the feminist theory written since the revival of political theory in 1971,

I consider that the more recent debate on ideal and non-ideal theory had made feminist theorist to move from the third to the forth project and that their reform advocated shifting from ideal to non-ideal theorizing.

The new view on non-ideal theory: abstraction without idealization and group relations methodology

Within the debate on abstract liberalism and the method used by liberal political philosophers that endorsed Kantian contractualism, Onora O'Neill (1987) offers a new perspective on the methodology of drafting ethical principles. She rejects relativist principles because, in her opinion, in order for ethical principles to be "relevant to a wide range of agents or situations, they surely not merely *may* but *must* be abstract" (O'Neill 1987, 55). However, how one arrives to an abstract principle is not an easy task. O'Neill is the first theorist to stress that there exists a difference between abstraction and idealization, and that in theories that issue principles there is a tendency to confuse the two concepts. For O'Neill, abstraction represents what is omitted, whereas idealization is what is added:

"The target of this line of criticism is not, however, just abstraction. The objection is not that much (too much) that is true of human agents is omitted in some accounts of agents, but that much (too much) that is false of human agents is added. Descriptions of agents in much post-enlightenment ethical and political theory are often idealized: they are satisfied only by hypothetical agents whose cognitive and volitional capacities human beings lack." (O'Neill 1987, 56).

This distinction helps also clarifying the feminist dilemma of choosing between abstract (and possibly gender blind in a patriarchal world) and relativized (that could reinforce patriarchal traditions in communities) theories of justice. Following O'Neill (2000, 163), my position is against a relativized theory of justice, because relativized accounts of justice end up legitimating gender-based forms of injustice (e.g. vulnerability), while idealized accounts of justice end up only ignoring them. The latter is easier to remedy than a justification of injustice. But what are, more precisely, the 'idealized accounts of justice'? For O'Neill (2000, 151), principles of abstract liberalism are in fact idealized principles as they add inexistent characteristics to agents such as rational choice in the context of perfect information, independence, etc. Of concern is not only that the liberal theories of justice that issue principles rely on false (added) assumptions on the human nature, but that this idealization leads to injustice by favoring certain social groups: "idealizations may privilege certain sorts of human agent and life and certain sorts of society by covertly presenting (enhanced versions of) their specific characteristics as true to all human action and life." (O'Neill 2000, 152). For O'Neill this is "idealization masquerading as abstraction" which leads to the "exclusion of those who do not match a certain ideal" (O'Neill 2000, 152). As such, idealization

used instead or in addition to abstraction leads to principles of justice that apply only to some individuals or to certain groups of individuals, and this perpetuates systemic injustice. This is also the case of Rawls' *Theory of Justice* and of most of the post-rawlsian theories of justice: in most theories of justice the focus on distributive patterns, where principles are issued in a thought experiment where all imagined agents are in fact idealized, from a white male biased perspective. For O'Neill, Rawls' constructivism is not an abstract one, but rather an idealized constructivism given the characteristics of the agents presented in the original position. This means Rawls' principles are not applicable to everyone and they can be rejected on this basis: "principles that cannot be acted on by all must be rejected by any plurality for whom the problem of justice arises" (O'Neill in Watkins and Fitzpatrick 2002, 352).

Given the distinction clarified by O'Neill (1987) between abstraction and idealization, one would wonder whether pure, 'correct' abstraction would not remedy ideal theorizing, and whether we could reaffirm the priority and necessity of ideal theory. Even if strict abstraction would be employed in ideal theory – take for example Rawls' Theory of Justice – it would still have deficiencies when it comes to certain group-based injustices such as racial and gender injustice. As Schwartzman (2006a, 2006b) highlighted, if abstraction is bracketing and omission of information, then we should also investigate who does this omission. In the classical sense, abstraction as method says nothing about deliberation or discourse, and thus allows the monological version of establishing what is relevant for justice.

The question that arises next is: given that relativized accounts of justice are not truly serving justice as they reinforce the status quo, and also given that abstract accounts of justice are in most of the cases abstract accounts of justice for idealized individuals, how should we arrive to a truly just account of justice? O'Neil (2000) opens the discussion of a new methodological pursuit, which has been picked up by feminists like Elizabeth Anderson (2010) and Iris Marion Young. For O'Neill (2000, 145), abstraction, even in its strict sense, is not enough as its principles will be too indeterminate. She proposes to link abstract principles with context-sensitive judgement of cases. This would better serve to draft principles that also serve gender and race justice, as abstract liberalism actually is not abstract, but rather male biased as it "almost always idealizes specific conceptions for example of human agents, of rationality, of family relations or of national sovereignty, which are often admired and are more (nearly) feasible for men rather than for women, and for developed rather than developing societies" (O'Neill 2000, 145).

Continuing O'Neill's (2000) proposal of linking abstract principles with particular cases, Elizabeth Anderson (2010) develops a new account of the methodology of non-ideal theory. She sees non-ideal theory as independent from ideal theory, and not a derivative of the latter. Anderson rejects ideal theory due to the fact

that she considers it “epistemically disabling, because its abstractions prevent thinkers from noticing certain forms of injustice” (Hertzberg 2014, 371) such as racial and gender injustice. As such, I would stress that Anderson is definitely also against transitional type of non-ideal theory that follows the implementation of principles identifies within ideal theory. Her argument against ideal theory is widely held among many feminists and critical race theorists (e.g. Mills 2005; Marion Young 1990, 2000).

Anderson, like Sen, considers that is better to address current problems instead of theorizing on the principles and ideas of a perfectly just society. Moreover, and here lies the novelty difference from Sen’s comparative justice, Anderson (2010, 3) proposes to link social sciences research with political philosophy. Further on, the connection between the two fields steps away from the classical domination of the normative political philosophy on the social sciences (in opposition with Rawls for whom ideal theory had lexical priority and would dictate its principles for implementation by social scientists). Anderson’s (2010, 3) methodological proposal is to “start political philosophy from non-ideal theory – from a diagnosis of injustices of our actual world, rather than from a picture of an ideal world”. She identifies three main reasons that would support this methodology: first, principles would be tailored on the real “motivational and cognitive capacity of human being”, or, in other words, this would respect Rousseau’s claim of taking people as they are and laws as they can be. Second, the topics addressed by the principles and ideas of justice would diversify and would look into the problems of the real world, which are not limited to distributive theories, but also include racial, gender and global injustice. Third, while “starting from ideal theory may prevent us from recognizing injustices in our non-ideal world” (Anderson 2010, 5), starting from non-ideal theory avoids this risk as it identifies group-based injustices like racism or sexism. Anderson thinks (and I agree) that principles that are color-blind are also injustice blind: “color-blind policy contradicts the color-blind ideal” (Hertzberg 2014, 374). Thus, starting from identifying racial and gender particular injustices serves reaching an abstract, neutral ideal.

The second particular proposal issued by Anderson on her method of doing non-ideal theory regards the role of ideals: after identifying forms of injustice, she proposed searching for ideals as a solution of reaching justice, and then testing them through the help of social science: ideals “function as hypotheses, to be tested in experience” (Anderson 2010, 6).

In my interpretation, Anderson does not propose simply to start observing the injustices of our actual world and to remedy them gradually. This would be Sen’s comparative justice. Instead, Anderson starts from non-ideal theory (so from our real world), but it then advances towards an ideal. Her ideal is not a relativized one, but rather an abstract ideal. The difference is she proposes to start from non-ideal (so, from stances of injustice) to the ideal abstract principles that would be inclusive and applicable to all members of the society. The transition to

the particular non-ideal to general ideal is made by using the method of strict abstraction, in O'Neill's terms, as this would mean passing from the particular to the general. As such, her method avoids the monological, isolated manner of issuing normative principles, which is the main approach in Rawlsian ideal theory. Even if the method of abstraction could be exercised by philosophers in solitude and detached from reality, testing the abstract ideal would require a democratic, deliberative, pluralist approach. Moreover, ideals are revisable, they are not fixed or imposed by an autocratic philosopher.

The fact Anderson's method allows revising its identified ideals comes in agreement with another perspective shared by another feminist political theorist, Alison Jaggar (2015) who considers that the method of "reversibility" is essential for ideal theorizing as it allows, through discourse, to "to adopt the perspectives of all others in the balancing of interests" (Habermas in Jaggar 2015, 114). Like Anderson, Jaggar also proposes to start from non-ideal when drafting principles of justice: "to maximize the probability of developing social ideals that are realistic in the sense of being contextually relevant, morally adequate, and practically feasible, I contend that philosophers should utilize a nonideal or naturalized approach to moral reasoning" (Jaggar 2015, 119). Jaggar (2015, 119-123) identifies six conditions in order for non-ideal reasoning to design a realistic utopia: (1) reasoning towards justification (RTJ) must be intersubjective and not monological, which means the method of discourse ethics is more suitable; (2) RTJ must be empirical not hypothetical; (3) RTJ must be inclusive; (4) RTJ must be fitted to specific contexts; (5) RTJ cannot be entrusted to philosophers alone; (6) RTJ must be reflexive.

A third account on the alternative to the 'classical' methodology employed by liberal theories of justice is given by Lisa Schwartzman (2006a). In *Challenging Liberalism. Feminism as Political Critique*, she dismisses liberalism on the basis of the two main methods it uses: individualism and abstraction. For Schwartzman (2006a), both methods undermine feminist goals. As such, she proposes to replace the two methods with their opposites: replace individualism with group relations, and replace abstraction with the non-ideal standpoint (Anderson 2009, 130). The difference between Anderson and Schwartzman is that Anderson does not reject abstraction as the method (abstraction in O'Neill's terms), but rather how (as idealization) and when (in thought experiments of ideal theory) abstraction is used. Anderson's methodology starts from non-ideal various stances of injustice and then drafts general principles that are inclusive. Schwartzman (2006b, 570), on the other hand, rejects abstraction (even O'Neill's version): if abstraction is bracketing or selective omission, then we are left with the issue of "what should be bracketed and what should be included". To this, I would also add the question of 'who' brackets (Schwartzman (2006a, 81). For this reason, Schwartzman (2006b, 586) considers that O'Neill falls victim of her own criticism.

Following MacKinnon, Schwartzman's (2006a, 85-86) main argument is against abstraction as methodology because, like many feminist legal theorists, she considers that supposed abstract theory actually assume a male norm. Take, for example, the state's intervention in domestic violence and the laws drafted for this purpose. The analytical method proposed by Rawls is vehemently rejected by Schwartzman (2006a) as for him power dynamics belong to the background culture and are not treated as political, hence not being addressed in the theory of justice. For feminists, the topics of power and oppression is essential in theorizing about justice. The incompatibility of mere abstraction and feminism is given, according to Schwartzman (2006a, 77), by the fact that abstraction makes oppression invisible and that attempts of abstracting endorse the dominant ideologies or the status quo. It is worth mentioning also that the detailed feminist critique addressed to abstract liberalism does not consider the latter too abstract, but rather "too male biased" (Schwartzman 2006a, 99). The same argument is used by critical race theorists, like Charles Mills (2005, 2014), who consider that "abstractions of ideal theory are not innocent" (Mills 2005, 181) and that "postraciality thus merges into preraciality, which is araciality, which is the abstract universal" (Mills 2014, 32).

As Babbitt underlines, abstract definitions of justice may take for granted and reinforce racist, sexist and classist power structures: "if sexist assumptions are embedded in the social context of meanings and values, they are not understood as sexist assumptions; rather, they are norms" (Babbitt 1996 in Schwartzman 2006b, 577-9). For Babbitt, abstraction might reinforce the status quo and this why ideals are necessary. However, ideals must also abstain from being biased or designed for idealized agents.

The second methodological change proposed by Schwartzman (2006a) is to shift the focus from individualism to group relations. This is a recurrent proposal in many feminist accounts of how one should do political philosophy and how to draft theories of justice.

Independent non-ideal theory and critical theory.

As I have shown above, both Anderson (2009, 2010) and Schwartzman (2006a, 2006b) agree on the importance of two main elements:

Non-ideal theory as a method that replaces ideal theory, but which seeks to treat the causes of injustice, not just to remedy them: "non-ideal theory [...] works toward diagnosis of the problem and then to practical solutions, is common to Marxism, Frankfurt school critical theory, and many versions of feminism" (Anderson 2009, 138).

Group relations methodology, because: “focusing on individuals as individuals may prevent one from noticing that often people are not regarded by society as mere ‘individuals’ and that various forms of injustice, power, and privilege stem from socially created group identities” (Schwartzman 2009, 181).

Both ideal theorizing and abstraction are essential for contract theory, and this is the reason what contract theory fails to represent the standpoint of oppressed groups (Anderson 2009, 136). The difference lies in the fact that for Schwartzman, liberalism is tied to the distributional account, whereas Anderson considers we can envisage a form of relational liberalism (Schwartzman 2009, 180).

As mentioned in the first part of the paper, Anderson’s proposal of starting with non-ideal theory (meaning with the lived experiences of various social groups in order to identify systemic forms of injustice) resembles similar proposals advocated by feminists who reject liberalism (e.g. MacKinnon, Schwartzman). In my view, this suggests that the methodology of non-ideal theory doesn’t have a fixed background political ideology: it is not a liberal form of theorizing. Moreover, this bottom-up approach of theorizing is shared also by Iris Marion Young (1990, 2000), but she calls her method critical theory. Similarly with Anderson, Jaggar, Schwartzman, and in contrast with Rawls’ methodology, Young proposes an inductive method, a bottom-up approach that would investigate the particular in order to arrive to general principles (Jaggar 2009, 97-98). I agree with Jaggar (2009, 97), who states that even if “Young does not explicitly characterize her method as ‘nonideal’, it diverges from all three ideal aspects of Rawls’s method”. In other words, Marion Young is also a feminist non-ideal theorist, even if she did not explicitly endorse the terminology.

In order to see if we can equate independent non-ideal theory with critical theory, it is worth mentioning how Young sees her methodology:

“critical theory does not derive normative principles and ideals from philosophical premises about morality, human nature, or the good life. Instead, the method of critical theory [...] reflects on existing social relations...” Young (2000, 10).

The definition above allows normativity to be part of her theory, but only by avoiding idealizations about social norms and human nature; social relations are starting point for her methodology. As such, according to Young’s description of critical theory, her method rejects both abstraction and idealized individualism, just like independent non-ideal theory. Somewhere else, she drafts critical theory as follows:

“critical theory is a normative reflection that is historically and socially contextualized. Critical theory rejects as illusory the effort to construct a universal normative system insulated from a particular society. [...] Without social theory, normative reflection is abstract, empty and unable to guide criticism with a

practical interest in emancipation. [...] Critical theory presumes that the normative ideals used to criticize a society are rooted in experience of and reflection on that very society, and that norms can come from nowhere else. [...] Critical theory is a mode of discourse which projects normative possibilities unrealized” (Young 1990, 5-6).

This definition also embraces the non-ideal rejection of abstract universalism issued from the white male philosophical perspective. As such, neither in critical theory, nor in non-ideal theory can ideals be imposed externally by a autocratic authority to arrives at these ideals in monological circumstances.

Given the above, I propose to integrate critical theory in the spectrum of possible non-ideal theories, as they both share the same approach on how to do political philosophy, and they both address stances of injustice faced by oppressed groups.

Equating critical theory with non-ideal theory generates a new dilemma: non-ideal theory appeared as a topic and method in the framework of liberal theories of justice; critical theory belongs to continental political philosophy. The following question arises: is it still worth to apply the analytical – continental distinction in political theory? The continental political philosophy has, among others, the following particular traits: (1) it is focused on the formation of power and domination; (2) has a historical oriented outlook (and hence identifies pervasive injustice); (3) “treats practices of political reasoning as embedded in broader social, economic and cultural structures” (Owen 2016, 179); (4) combines social theory with political philosophy. Given that all these traits can be also identified in the structure of independent non-ideal theory, is there an essential difference between non-ideal theory and continental political philosophy?

I disagree with Owen’s (2016, 185) conclusion that non-ideal theory is analytical (he uses Anderson’s theory as example) and that Young’s critical theory is different from continental political philosophy. In my perspective, non-ideal theory does not strictly belong to analytical political theory, neither to liberalism. While Owen (2016, 185) considers that Young’s method is a hybrid (he calls it “activist political theory”) that can bridge analytical and political philosophy, I consider that there is no need to build the bridge between the two, as non-ideal theory already uses the methods advocated by continental political philosophy.

References

Anderson, Elizabeth. 2010. *The Imperative of Integration*. Princeton: Princeton University Press.

- Anderson, Elizabeth. 2009. "Toward a non-ideal, relational methodology for political philosophy: comments on Schwartzman's *Challenging Liberalism*." *Hypatia*, 24: 130-145.
- Estlund, David, 2014. "Utopophobia." *Philosophy & Public Affairs*, 42: 113-134.
- Gilabert, Pablo, and Lawford-Smith, Holly. 2012. "Political Feasibility: A Conceptual Exploration." *Political Studies*, 60 (4): 809-825
- Gilabert, Pablo. 2012. "Comparative Assessments of Justice, Political Feasibility, and Ideal Theory." *Ethical Theory and Moral Practice*, 15: 39-56.
- Jaggar, Alison. 2015. "Ideal and nonideal reasoning in educational theory." *Educational Theory*, 65: 383-408.
- Jaggar, Alison. 2009. "L'imagination au pouvoir: Comparing John Rawls's method of ideal theory with Iris Marion Young's method of critical theory", In *Feminist Ethics and Social and Political Philosophy: Theorizing the Non-Ideal*, Tessman L. (eds). Dordrecht: Springer, 95-102.
- Hertzberg, Benjamin. 2014. "Ideal and non-ideal theory in Elizabeth Anderson's *The Imperative of Integration*." *Political Studies Review*, 12: 369-375.
- Kittay, Eva Feder. 2001. "When Caring is Just and Justice is Caring: Justice and Mental Retardation." *Public Culture*, 13: 557-579.
- MacKinnon, Catharine. 1989. *Toward a feminist theory of the state*. Cambridge: Harvard University Press.
- Mills, Charles. 2005. "'Ideal Theory' as Ideology." *Hypatia*, 20: 165-183.
- Mills, Charles. 2014. "White time. The chronic injustice of ideal theory." *Du Bois Review*, 11: 27-42.
- Okin, Susan Moller. 1989. *Justice, Gender, and the Family*. New York: Basic Books.
- O'Neill, Onora. 1987. "Abstraction, Idealization and Ideology in Ethics." *Royal Institute of Philosophy Lecture Series*, 22, 55-69.
- O'Neill, Onora. 2000. "Gender, justice and international boundaries." In *Political bounds of justice*, pp. 143-167.
- Owen, David. 2016. "Reasons and practices of reasoning: On the analytic/continental distinction in political philosophy." *European Journal of Political Theory*, 15: 172-188.
- Rawls, John. 2005. *A Theory of Justice*. Cambridge: The Belknap Press of Harvard University Press.
- Robeyns, Ingrid. 2008. "Ideal Theory in Theory and Practice." *Social Theory and Practice*, 34: 341-362.

- Schwartzman, Lisa. 2006a. *Challenging Liberalism. Feminism as Political Critique*. University Park: Penn State University Press.
- Schwartzman, Lisa. 2006b. "Abstraction, idealization, and oppression", *Metaphilosophy*, 37: 565-586.
- Schwartzman, Lisa. 2009. "Non-ideal theorizing, social groups, and knowledge of oppression: a response." *Hypatia*, 24: 177-188.
- Sen, Amartya. 2006. "What do we need from a theory of justice?" *Journal of Philosophy*, 103: 215-238.
- Sen, Amartya. 2009. *The Idea of Justice*. Cambridge: The Belknap Press of Harvard University Press.
- Simmons, A. John. 2010. "Ideal and Nonideal Theory." *Philosophy and Public Affairs*, 38: 5-36.
- Srinivasan, Amia. 2018. "Feminism and Metaethics." In *The Routledge Handbook of Metaethics*, (McPherson, Tristan and Plunkett, David, Eds.), 595-608.
- Watkins, Eric, and Fitzpatrick, William. 2002. "O'Neill and Korsgaard on the Construction of Normativity." *The Journal of Value Inquiry*, 36: 349–367.
- Watson, Lori. 2013. "Toward a feminist theory of justice: political liberalism and feminist method." *Tulsa L. Rev.*, 35: 35-44.
- Young, Iris Marion. 1990. *Justice and the Politics of Difference*. Princeton: Princeton University Press.
- Young, Iris Marion. 2000. *Inclusion and Democracy*. Oxford: Oxford University Press
- Zerili, Linda. 2008. "Feminist theory and the canon of political thought." In *The Oxford Handbook of Political Theory*, 106-124.

4. Keeping Non-ideal Theory Real

WORKING PAPER by David Jenkins (University of Warwick)

In recent years, much of contemporary political theory has been increasingly criticised for abstracting from political, social and economic reality. The discipline, the criticism continues, is infected by a basic category mistake about what politics is and what political theory should do and say about it. The positive content of these criticisms is the suggestion that we move away, as theorists, from ideal theory and toward the non-ideal. Others critics, while still sanguine about the uses ideal theory performs, see the problem as one of a division of labour. Currently too much time is spent with ideal theory and not enough engaged with complex and pressing, real world problems.

There are, broadly, two ways to define the conditions that constitute non-ideal theory: One derives from a separate account of *injustice*, the other from a description of the deviation from norms specified by ideal theories of justice. In this essay, I will have little to say about the relative merits of the first of these forms of theorising. Nevertheless, it is implicit from my argument that one of the merits of such theorising is it evades the problems here identified. My critique is directed at a second style of describing non-ideal conditions in terms of agents' partial or non-compliance with the demands of justice. It is this that is adjudged to define a problem – more often that not *the* problem – we confront when we confront non-ideal conditions. I find the explanatory power of this move not only underwhelming but obfuscating of the *real* conditions underlying non-ideal situations.

Under ideal conditions, each citizen is assumed to fully comply with the demands of justice and to *know* that all other citizens are similarly motivated to comply. This has consequences for the institutions that structure the ideal world, as well as the personalities of the individuals who will have their obligations, entitlements and certain parts of their behaviour constrained by these institutions. A major part of the procedures associated with ideal theory is dedicated to defining the reasonable, shared costs individuals will have to accept as part of an assumed commitment to justice. I defend this function against the charge that it is an idealisation and thus an illegitimate component of ideal theorising.

A crucial part of the shift from ideal to non-ideal conditions is recognising that whereas assuming full-compliance produces mutually acceptable cooperation, in non-ideal conditions the existence of any minimally

mutually beneficial cooperation is necessarily accompanied by an additional condition of ongoing complicity with injustice. It is this complicity with injustice that should, I argue, replace any talk of partial compliance: It is the fact that one is always, irrespective of motives, inescapably involved in structurally embedded injustices, which produces obligations that should then motivate one's behaviour.

I shall argue first, that the option of non-complicity is more or less impossible, second that there exist substantial differences in the benefits differently positioned people receive as a consequence of complicity, and third that there are various degrees of culpability that track both the benefits one is able to receive from one's complicity and the power one enjoys *vis-à-vis* the practices and institutions definitive of (unjust) 'basic structures'. By approaching partial-compliance with this more nuanced understanding of complicity in mind, we are able to break that category down, revealing important differences between different varieties of partial-compliance. One category in particular, a category I label "constitutive subversion", is of especial relevance. In conclusion, I suggest that understanding the nature of this particular category reveals something of a tragedy characterising non-ideal conditions.

I: Starting Points

Before proceeding I believe it is necessary to make a preliminary note with regards to vocabulary: First, I have always found the use of 'non-ideal' to describe the real world a little euphemistic. In a world with some 30 million slaves, where three of the most globally traded 'goods' are people, drugs and guns, and where we are daily confronted with stories of degradation, exploitation and systemic corruption, non-ideal sounds pretty inept as a description. However, it is ubiquitous across the literature and so I will retain it.

In less formal terms then, non-ideal theory is theory concerned with the world as it is. The lack of clarity in these terms has led to some confusion: For example, Rawls' treatment of civil disobedience – cited as one of the few instances when non-ideal conditions occupy his attention – is supposed to be aimed at societies of 'near

justice'.¹⁵ However, when we consider where civil disobedience has historically taken place – the American South for instance – it would be an enormous – perhaps even offensive – stretch to try and call these societies anywhere near approximate to justice. It is not clear then, whether the relevance of this description of disobedience is being used *vis-à-vis* a more or less idealised non-ideal world, i.e. a world we have yet to see, but which still fails to hit the heights of the truly ideal. However, if there is a useful distinction between the real world and non-ideal conditions, I do not consider it in this paper.¹⁶

Of course this confusion points to the complications non-ideal theorists confront once they start allowing the 'world' into their theorising. Ideal theory avoids this issue by bracketing huge parts of that world.¹⁷ This is not to suggest that ideal theorists have no concern or use for facts *per se*, but those facts are made sufficiently hygienic and separate from facts as they exist 'out there', in a world characterised by profound injustices. Rawls' 'circumstances of justice' for example are intended to track the conditions of modern, ostensibly democratic societies but are abstractly defined and do not assume anything like the inequality, oppression and violence that permeates reality. But for the non-ideal theorist it is precisely this issue of hygiene that is being jettisoned in favour of engagement with the facts of the world.

David Foster Wallace's description of the difference between fiction and non-fiction writing is usefully analogous to this ideal/non-ideal distinction:

¹⁵ John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard Belknap Press, 1999) section 57.

¹⁶ I am, however, in agreement with Matt Sleat, against Laura Valentini, that realism offers a different approach to political theory that is not captured by a continuum that starts from the ethereally abstract, and gradually lets in more and more reality, such that realism is equivalent to a form of non-ideal theory, one that allows in substantial amounts of reality. Realism's approach and focuses are fundamentally different to what has been described as non-ideal theory. Given this, however, it seems non-ideal theory could call itself *realistic* without thereby becoming 'realist'. See respectively Matt Sleat, 'Realism, Liberalism and Non-ideal Theory Or, Are there Two Ways to do Realistic Political Theory?', *Political Studies*, 64/1 (2016) pp 27 – 41. Laura Valentini, 'Ideal vs. Non-ideal Theory: A Conceptual Map', *Philosophy Compass*, 7/9 (2012) pp 654 – 664.

¹⁷ Lois McNay, *The Misguided Search for the Political* (Oxford: Oxford University Press, 2014) p 6.

“Writing-wise, fiction is scarier, but nonfiction is harder—because nonfiction’s based in reality, and today’s felt reality is overwhelmingly, circuit-blowingly huge and complex. Whereas fiction comes out of nothing. Actually, so wait: the truth is that both genres are scary; both feel like they’re executed on tightropes, over abysses—it’s the abysses that are different. Fiction’s abyss is silence, nada. Whereas nonfiction’s abyss is Total Noise, the seething static of every particular thing and experience, and one’s total freedom of infinite choice about what to choose to attend to and represent and connect, and how, and why, etc”.¹⁸

In truth, the ideal theorists’ abyss is less ‘scary’ than the fiction writers. There is a more or less established subject matter – the ideal political community – and so there are the various facts the theorist will want to condition the output. At a minimum, conditions of mortality, assumptions about moderate scarcity and limited altruism. But the non-ideal theorist does seem to face this dilemma of noise and connection: The political (social, economic) world is a hugely complex, dynamic system of many different kinds of interactions and relationships. In all that (relevant-to-theorists) noise, the question becomes: where to begin?

Amartya Sen offers one way of making sense out of, and providing guidance *through*, that noise. Sen’s methodological approach involves an explicit rejection of the need for ideal theory to provide normative guidance, rejecting Rawls’ argument that the more fundamental grasp of pressing problems requires prior theorising in the ideal vein.¹⁹ Though the transcendental method employed by ideal theorists might offer an interesting intellectual exercise, it can tell us nothing – so Sen argues – about the comparative merits of institutional arrangements that might reduce the manifold injustices we confront, and are feasible from within contemporary circumstances. The unique identification of the just society is unnecessary as a starting point and

¹⁸ David Foster Wallace, *Both Flesh and Not* (London: Penguin, 2013) p 302.

¹⁹ John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), p. 285.

has absorbed far too much intellectual effort. Consequently, theory should begin with understanding the effects of injustice and then contribute to our understandings of how to remedy them.²⁰

Sen aims to generate normative claims and prescriptions from a prior assessment of the realities of the world. Ideal theory, on the other hand, is, as Iris Marion Young put it, ‘detemporalized’: “nothing that comes before it, (nor) future events will ... affect its truth or relevance to social life”.²¹ Its offerings are, according to Sen, akin to the “grand revolutionary’s ‘one shot handbook’” in which a plan is drawn up to be then implemented by various well-placed agents.²² This is not only regarded as methodologically suspect, but as abstracting from the constitutive material of normative theory, which should be actually existing problems in the world.

Sen’s is a powerful critique of ideal theory. What it demonstrates is an alternative starting point from which non-ideal theory can begin. For reasons of space, I assume that Sen is wrong in his suggestion that ideal theory is *necessarily* the wrong way to think about justice. In so doing, I also stake a claim against realists that has to go undefended.²³ My argument is directed against the way in which the move from ideal to nonideal theory has been theorised (in the main), but is not against the move itself. In other words, the argument I develop here

²⁰ Amartya Sen, *The Idea of Justice* (Cambridge, Mass.: Belknap University Press, 2009) pp. 87 – 113. See also David Wiens, ‘Prescribing Institutions without Ideal Theory’, *The Journal of Political Philosophy*, 20/1 (2012) p 45. Elizabeth Anderson’s understanding of non-ideal theory, and the thorough interdisciplinary work that underlines her approach, is also motivated by a different set of concerns, and no mention is made of full or partial or non-compliance. See Elizabeth Anderson, *The Imperative of Integration* (Princeton: Princeton University Press, 2010). 3 – 7. Tommie Shelby, *Dark Ghettos* (Boston: Harvard University Press, 2017) 11 – 12 offers another take, where non-ideal theory is ultimately derivable from ideal theory, but we have enough of the latter to be able to get on with the former.

²¹ Iris Marion Young, *Justice and the Politics of Difference* (Chichester: Princeton University Press, 1990) p 4.

²² Sen, *The Idea of Justice*, p 100. See also Bernard Williams’ description of political moralism’s tendency to produce ‘enactment models’. Bernard Williams, *In The Beginning was the Deed* (Oxford: Oxford University Press, 2005) p 1.

²³ Given the title of the paper, it should be clear that I do think nonideal theory should attempt to be more realistic in its descriptions of the real world. However, I do not assume any of the realist assumptions with respects to the ‘proper’ understanding of politics.

allows that there might be value in at least certain kinds of ideal theorising and these can help how we do nonideal theory.

The method of doing non-ideal theory under analysis here is that which takes its terms from ideal theory: Or, more accurately, that describes the non-ideal in terms of the distance between the real world and one that has been ideally conceived. For example, George Sher formulates non-ideal conditions as “lapses from the ideal (which offer) independent occasions for normative reflection”.²⁴ The non-ideal is thus an independent field of enquiry, but not on its own terms: Instead, those terms are derived from the failure of *ideal* terms to hold in the world at large. Its independence is thus generated when that ‘lapse’ becomes sufficiently radical or extensive.

As a result, in order to grasp the constituent parts of the relevant non-ideal theory there is a prior need to establish what we mean by the ideal. This does not need to be an extensive fleshing out of the ideal, but only the general terms of what we mean by ideal. This will set the standards from which the real world is to be considered a deviation, and also identify the relevant agents, the capacities and motivations of those agents, the costs they will be expected to bear, as well as the institutional framework within which they will pursue their various ends. The non-ideal world involves deviations across these various constitutive parts of the ideal. A central aspect of that deviation is the lapse from an ideal of ‘full-compliance’ to a reality of ‘partial-compliance’.²⁵

Ila: Functions of Compliance

²⁴George Sher, *Approximate Justice* (Oxford: Rowman & Littlefield Publishers, 1997) p 1.

²⁵ The other ‘lapse’ commonly mentioned is the possibility for there to be a move from the moderate scarcity assumed by ideal theory, to immoderate scarcity. However, much less is said about this, perhaps because scarcity in the slums of Johannesburg does not feel like an accurate description of what is going on in the presence of the superabundance just across the road – and, of course, across the militarily protected fence.

Zofia Stemplowska provides a broad account of the procedures associated with ideal theory as a technique of political theorising.²⁶ Theory, in a general way, is composed of inputs, outputs and the rules regulating the derivation of outputs from inputs. Specifically *normative* theory will contain normative principles as both inputs and outputs – it is this which makes them normative. An important difference between non-ideal and ideal normative theory focuses on the presence or absence of a particular input: Full-compliance. So while there is not a single position regarding ideal theory's precise function, the assumption of full-compliance exists across all of them.

Over and above this ubiquity however, is the *centrality* of full-compliance to all of ideal theory's functions. For example, Stemplowska defines ideal theory in terms of the “non-AD recommendations” it makes – recommendations that are neither achievable nor desirable as things stand.²⁷ Assuming full-compliance as grounds for these kinds of recommendations, Stemplowska argues, helps us get a clearer understanding of our ultimate values and principles.²⁸ Ingrid Robeyns suggests full-compliance can be used as a way to show us the ‘paradise islands’ or ultimate goals we should try and collectively approach.²⁹ Similarly, Robert Jubb describes full-compliance as a means by which to understand both perfection, which in turn informs our sense both of our distance from that perfection and complicity with injustice.³⁰ Finally, John Simmons argues that assuming full-compliance is necessary to inform the directions of improvements to be made by non-ideal theory, a reference point by which normative questions in the here and now can be grounded.³¹

²⁶Zofia Stemplowska, ‘What’s ideal about ideal theory?’, *Social Theory and Practice* 34/3 (2008) pp 323 – 324.

²⁷*Ibid.* pp 324 – 325.

²⁸*ibid.* p 330.

²⁹Ingrid Robeyns, ‘Ideal Theory in Theory and Practice’, *Social Theory and Practice* 34/3 (2008) pp. 341 – 362.

³⁰Robert Jubb, ‘Tragedies of Non-ideal theory’, *The European Journal of Political Theory* 11/3 (2012) pp. 229 – 246.

³¹Simmons, ‘Ideal and Nonideal Theory’, *Philosophy and Public Affairs* 38/1 (2010) pp. 34 – 36.

Full-compliance is really the assumption that *enough* people adhere to the principles that well-order a society so that the costs associated with sustaining just institutions are shared in a way that can be justified. The qualifier “enough” is crucial. I frame this slightly differently to Hamlin and Stemplowska who argue full-compliance requires that “*almost* everyone does *almost* everything that the normative content of that theory demands of them”.³² If the relevant threshold is one of sufficiency to guarantee the operating of just institutions and practices, this should be made clear. Such sufficiency *might* require almost everyone to do almost everything demanded of them, but the issue remains one of sufficiency not of near completion. The presence of free-riders, for example, is only a problem if they undermine possibilities of justice. Even criminal behaviour can be accepted to some degree without it thereby becoming problematic.³³ There is a threshold below which presence of such action might remain a nuisance, a reason for censure and even punishment, but not a concern from the perspective of sustaining the relevant institutions and practices. If we want ‘always and everywhere full compliance’, this is a different and lesser utopian demand than that of ‘sufficiently full compliance’.

We assume strict compliance because, as Simmons quite rightly suggests, it is difficult to theorise instances of *problematic* non-compliance in the abstract. And so he asks: “Exactly what form does the assumed noncompliance take, for instance, and how does that form of noncompliance itself influence the behavior and the expectations (etcetera) of others? What will we understand to be the “normal” forms of noncompliance, given that the behaviors or the institutional structures that count as noncompliant with different, competing principles

³² Alan Hamlin and Zofia Stemplowska, ‘Theory, Ideal Theory and the Theory of Ideals’, *Political Studies Review* 10 (2012) p 49. My italics.

³³ Nils Christie, *A Suitable Amount of Crime* (London: Routledge, 2004). Indeed, Norway – Christie’s home turf – would be just such a working example of the possibility for initiating substantively just social practices alongside humane systems of rehabilitation set up to handle non-compliance. We can have a just basic structure and a broadly just society with some murder in it.

of justice must be themselves quite different?”³⁴ But once we move to the real world, only those instances of behaviours, practices and actions that upset the justice of the basic-structure should trouble us.

Drawing on Onora O’Neill’s distinction, assuming full-compliance is taken to be an idealisation, not an example of necessary (and more innocent) abstraction. O’Neill defines an abstraction as the practice of “bracketing, but not of denying, predicates that are true of the matter under discussion”. Idealisations on the other hand are predicates that are “enhanced, idealised” that “are false of the case in hand”.³⁵ From O’Neill’s perspective abstraction is an aspect of theory we can legitimately embrace, a means of condensing that Total Noise Wallace describes above into its most relevant elements. But idealisations like full compliance are not so innocent. For example, adopting enhanced/idealised capacities of the person builds into the theory “unvindicated assumptions” that fail to test any consequent argument based on those assumptions “against the limitations of the human condition and the constraints of human knowledge”.³⁶

This distinction, as identified by Hamlin and Stemplowska, is problematic, most especially in determining the criteria to be used when deciding whether a false assumption is vindicable or not?³⁷ Since the goalposts have already been moved from strict-compliance to, when applied to the real world, sufficient compliance, the alleged utopianism of the assumption seems less obvious. At the very least, the utopianism of the assumption takes on a different hue. In Brian Barry’s theory of ‘justice as impartiality’, the assumption of strict-compliance is useful in specifying the costs to be borne by individuals in a situation where everyone is doing what they should be doing *vis-à-vis* the demands of justice. Under the assumption of full-compliance, these costs are by their very nature considered reasonable. Barry’s theory uses the Scanlonian principle of “reasonable rejection” to ground the relevant criteria that will “determine what principles are consistent with the ‘strains of commitment’ test”, i.e.

³⁴ Simmons, ‘Ideal and Nonideal Theory’, p 8 – 9.

³⁵ Onora O’Neill, *Towards Justice and Virtue* (Cambridge: Cambridge University Press, 1996) p 40-41.

³⁶ *Ibid.* p 63

³⁷ Hamlin and Stemplowska, ‘Theory, Ideal Theory and the Theory of Ideals’, p 50.

which citizens can be reasonably obligated to accept. Here, everyone is assumed to be motivated by a desire to come to reasonable agreement with all relevant and participating others, when deciding on the principles by which they will be, collectively, governed.³⁸ If some demand could be reasonably rejected by anyone because of the extent of the costs imposed on them, then this counts decisively against the justification of that principle.

In ideal theory, strict-compliance precludes the very possibility for complicity in injustice. We assume strict-compliance in order to design institutions and practices within which the fair mutual involvement and cooperation between relevant agents is guaranteed. The partial-compliance with the demands of justice that is supposed to characterise non-ideal theory means that, although there remain possibilities for *degrees* of cooperation and fair mutual involvement, there is also that additional factor of being unavoidably complicit in injustice. A question marginalized but centrally relevant for non-ideal theory is thus what a person should do in light of his or her own complicity, as well as that of others, in the production and reproduction of those injustices.

A further consequence of the shift toward non-ideal conditions is that the possibility for ideal compliance in a non-ideal world is precluded: In a world structured by idealised assumptions of full-compliance, questions of what one should do about injustice cannot arise in any substantive way.³⁹ In the real world, on the other hand, there is no possibility to comply with a fundamentally and comprehensively just ‘basic structure’ that would release individuals from involvement with processes of injustice. The idea of compliance supposes that there is

³⁸ Brian Barry, *Justice as Impartiality* (Oxford: Clarendon Press, 1995) p 67. It is thus, to take a point made by Jacob Levy and David Schmidtz, not about us being James Madison’s angels, for whom government is unnecessary. It is, rather, about us *not* being Madison’s ‘beasts in men’s shape’ constituting the ‘giddy multitude’, an image of the majority of humanity which so severely influenced Madison’s ambivalence regarding democracy. It is telling that realists so often and so complacently draw on Madison’s elitist scorn to support their own arguments. Jacob Levy, ‘There is no such thing as ideal theory’, *Social Philosophy and Policy*, 33/1 – 2, 2016. 312–333; David Schmidtz, *Person, Polis, Planet* (Oxford: Oxford University Press, 2008) 8.

³⁹ This is one of the shortfalls in Murphy’s description of collective beneficence: Assuming strict-compliance to define non-ideal duties excludes complicity with injustice as a relevant concern for those who will act. See also Jubb, ‘Tragedies of Non-ideal Theory’, p 241.

something prior to the choices and decisions individuals make which can form the subject of those choices and decisions: Compliance is, we might say, a fundamentally bureaucratic act. It involves, that is, the knowledge that our performing some particular task or set of tasks will mean we act in ways that count as compliant: I do A and B, I have complied. People's not complying with the nonideal demands of justice, for example, by failing to commit to the long, messy and unpredictable work needed to create just institutions, tells us nothing about what they might do against a background of sufficiently just institutions.

Compliance in ideal conditions produces both assurances of others' compliance and the material, institutional *consequences* of those assurances which are then reiteratively reinforced by one's own decision to continue complying. However, when there is nothing at the institutional level that can be complied with in order to produce a substantively just situation, it does not make sense to hold the person accountable, or to blame her, for her not doing something she never had the option to do: In other words, to describe the non-ideal conditions that are the result of that failure as the *result* of partial compliance. Partial compliance might only be the consequence. Our complicity with mass injustice is not the result of an across the board failure in individual acts of compliance.

Ilb) Function of Partial Compliance

It should be clear then that I am not criticising the assumption of full-compliance in building *ideal* theories of justice. Nor, again, am I committed to Sen-style critiques of ideal theory. The problems I identify only occur when we move from what are intentionally ideal suppositions of full-compliance, to non-ideal assumptions of partial-compliance. So, where ideal theory imagines what can be accomplished when enough people are motivated to a

sufficient enough degree to do their part, non-ideal theory frames its normative concerns with the question as to what people ought to do in circumstances when significant numbers of others are *not* doing their part.⁴⁰

It is this shift in the terms of compliance that is supposed to define the central normative concern of non-ideal theory.⁴¹ The full-compliance of everyone assumed by ideal theory is thus broken down into everyone's partial or non-compliance: The input of everyone fully complying is replaced by the input of everyone only partially complying, and thus we arrive at the shape and shade of the real world. Only once these circumstances have been delineated is the question asked as to what, in the real world of partial compliance, individuals motivated by demands of justice should do. The normative guidance subsequent to that delineation is framed in terms of the "fair shares" of what one is obliged to do in a world where others are only ever partially compliant.⁴² In the non-ideal world, this fair share can be understood in very different ways, ranging from there being no difference to what one would owe in situations of strict-compliance, to incredibly demanding costs that would severely reduce a person's possibilities to do anything other than attend to the needs of the desperately impoverished.

Liam Murphy, as an exemplar of the first position, outlines a "collective principle of beneficence" – contrasted with utilitarianism's *optimizing* principle – in which "the demands on a complying person should not exceed what they would be under full-compliance with the principle".⁴³ To have it otherwise, to assume that individuals must

⁴⁰ Laura Valentini, 'Ideal vs Non-ideal Theory: A Conceptual Map', *Philosophy Compass* 7/9 (2012) p 655. Liam Murphy, *Moral Demands in Nonideal theory* (Oxford University Press, 2000, p 5; Andrea Sangiovanni, 'Normative Political Theory: A Flight from Reality?' in Duncan Bell (ed.) *Political Thought and International Relations* (Oxford: Oxford University Press, 2008) p 221.

⁴¹ The other refers to the 'historical, cultural and social and economic conditions'. This is boiled down to the having the technical sophistication of institutions in a 'democratic regime', and is marginalised in favour of focussing on issues of compliance. Stemplowska, 'Non-Ideal Theory', 287. See also John Rawls, *Justice as Fairness: A restatement*. (Cambridge MA: Harvard University Press). 47.

⁴² See for instance David Miller, *Justice for Earthlings* (Oxford: Oxford University Press, 2013) pp. 206 – 227.

⁴³ Murphy, *Moral Demands in Nonideal theory*, p 7.

pick up the slack as a consequence of others' failures, is to make excessive and unjust demands of them. Individual sacrifices might still need to be made, but by assuming others' compliance everyone is only ever doing their bit, and never somebody else's.

At the other extreme, Elizabeth Ashford has argued that from within Tim Scanlon's contractualist approach to morality, the demands that one faces in situations of minimal compliance and significant injustice are inescapably huge. The Scanlonian notion of 'reasonable rejection' implies that a persons' behaviour is morally acceptable if and only if it is not prohibited by a set of moral principles that no one could reasonably reject. What counts as reasonable rejection depends on the relative burden of costs that have to be justified to those who will be most burdened by them. In light of the extreme burdens suffered by some in current circumstances, it is hard to see, Ashford suggests, how any 'relatively affluent' individual could justify anything less than behaviour implied by Peter Singer's – admittedly and unabashedly – demanding principle of maximal beneficence.⁴⁴

More generally though, there are problems with any approach to the nonideal that describes the subject of such theory as the result of failures in terms of individuals' compliance, partial or non-compliance. David Schmidtz, for example, criticizes ideal theory because it assumes something that cannot be taken for granted in the real world: 'To say "ideally we would not have compliance problems" is like saying "ideally we would not need to drive defensively"'. This, he suggests, is a 'remark about a world whose problems, and therefore whose solutions—whose ideals—are not like ours.'⁴⁵ For Schmidtz, when we begin to theorise in political theory, failures of compliance are the fundamental real world fact we must confront. To start anywhere else is to fail to produce

⁴⁴ Elizabeth Ashford, 'The Demandingness of Scanlon's Contractualism', *Ethics* 113 (2003) pp. 273 – 293.

⁴⁵ David Schmidtz, 'A Realistic Political Ideal', *Social Philosophy and Policy* 33/1. 6

properly political theory.⁴⁶ Problems of compliance are problems that come with taking people as they are. We are fundamentally, it seems, creatures who balk against compliance.

For certain understandings of the connection between ideal and nonideal worlds, this understanding might ring true. For example, where principles of justice require too much in the way of ‘goodness’ from individuals, then this issues in demands of compliance that are always going to be out of reach.⁴⁷ We can locate these defects in the person or in the theory, but whatever we do, the problems that manifest as a consequence of this kind of non-compliance do not describe a non-ideal world approximate to the real world. Moving from challenging the assumption of full compliance in ideal situations, to positing the inevitability of partial compliance in any world we can hope to achieve, provides little of value when describing the world as we find it today.

III: The Real Nonideal

Describing nonideal contexts in terms of failures of compliance misses crucial distinctions within the group of people described as failing in their duties to comply. To concentrate on the fact – assuming it is, broadly, a fact – that everyone is only ever (at best) partially complying with the demands of justice, is to overlook important differences within the party of the accused. Attempting to describe the nonideal world in terms of deviations from an ideal of full compliance is to obfuscate the realities of that world. Most especially, it assumes an equality amongst partially complying agents which needs to be unpacked. By interrogating this assumption, recognising

⁴⁶ As much is accepted by Zofia Stemplowska when she describes people’s noncompliance as that which ‘blocks’ the ideal option, without disaggregating the enormous groups of people this throws together. Zofia Stemplowska, ‘Non-Ideal Theory’, Lippert-Rasmussen, Brownlee, Coady (eds.) *A Companion to Applied Philosophy* (Chichester: John Wiley and Sons, 2017). 287.

⁴⁷ Take this from David Estlund’s description of the problem: “People could be good, they just aren’t. Their failures are avoidable and blameworthy, but also entirely to be expected as a matter of fact. So far, there is no discernible defect in the theory, I believe. For all we have said, the standards to which it holds people might be sound and true. The fact that people won’t live up to them even though they could is a defect of the people, not of the theory.” David Estlund, *Democratic Authority* (Princeton: Princeton University Press, 2011) 264.

important differences between the agents who fill out nonideal worlds, reduces the importance of non-compliance in our descriptions of those worlds. However, this does not mean we should necessarily abandon concerns with compliance as an important feature of theorising about justice. Instead, I shall argue that recognising impossibilities for compliance in non-ideal conditions should serve to turn our attention to another couple of concepts – *complicity* and *culpability*.

It is first necessary to determine whether it is possible to describe individuals absent some mutual involvement with each other, as being partially- or non- compliant with the demands of justice: Before accusing someone of being partially compliant with the demands of justice, is it first necessary to set out some kind of *relationship* between the partially compliant? Or does this partial-compliance stand over and above such relationships? Does ideal theory, specifically the assumption of strict-compliance, imply natural duties of justice that emerge for any and every one, regardless of one's role in the production or reproduction of injustice?

In order to answer this question, we need to return to the ideal from which non-ideal partial-compliance describes a deviation. Full-compliance is assumed in the process of justifying institutions to those individuals who are going to be affected by the operations of those institutions. This supposes a relationship of ongoing, mutual involvement and cooperation at the basic level of ideal theory. This is then translated – in suitably lapsed form – to the realm of non-ideal theory. Both levels take as given some form of community and mutual involvement between the ideally compliant or nonideally partially-compliant parties.

Imagine a perfectly self-sufficient society that has never had any connection with the outside world. All around this little community lie mass injustices that are a consequence of problems located exclusively within and between the communities of that wider world. What duties do the members of such a society owe to people beyond their borders? Does anything like compliance with the demands of justice – full or otherwise – apply to this community? Regardless of the community's capacity to help these other people, in the absence of a separately defined duty to promote the justice of institutions wherever that might be possible, it seems wrong to suggest that this community's refusal to offer any assistance can be described as a failure to comply with the demands of justice – at least as derived from the operations and assumptions grounding ideal theory. Such a community would not be considered party to the assumptions of full-compliance from which the reality of partial-compliance is a deviation. The non-ideal world describes a situation where ongoing, mutual involvement has degenerated into, for example, relationships of oppression and domination.

Of course, in a globalised world economy with a convoluted history of mostly violent and exploitative interactions, any notion of hygienic self-sufficiency is anachronistic by at least several hundred years – especially for those residing in modern, capitalist societies. Additionally, any society that remains aloof from this history is

unlikely to be able to offer any substantive assistance whatever the status of their duties might be. But the distinction is nonetheless relevant for grounding assessment of partial-compliance in a prior, more or less substantive relationship, i.e. beyond shared membership of the same species. The issue of complicity thus emerges as part of what partial-compliance is intended to capture: In non-ideal conditions one is involved in processes and practices that produce both benefits and costs, the distributions of which are unjust. Involvement in these injustice-producing processes trigger both *obligations* regarding what one should do about that complicity, and *entitlements* in light of the burdens one assumes.

In reality, the distribution of burdens is not always easily reduced into the categories of victimhood and beneficiary. People can simultaneously benefit from the injustices under which others labour, contribute to those sufferings by their involvement in the practices and interactions that produce injustice and, finally, *themselves* suffer in various ways from the injustices that characterise the world.⁴⁸ Non-ideal conditions reap both (unequal) rewards and produce (unequal) burdens for most people most of the time.⁴⁹

For an example of these over-lapping aspects, consider the kinds of actions and interactions that go into the production and consumption of goods. How we produce and consume in the world is always from within an on-going set of practices which are connected in myriad ways with the performance of similar acts by others, most of whom will assume significant costs as a result of their participation. In the complex interactions definitive of this context, even the poor in advanced western states benefit in fairly obvious ways from the immiseration of others. The clothes they wear, electronics they use and food they eat are all the consequences of productive practices steeped in exploitation and coercion.⁵⁰ Going further down the supply chain: Workers in Chinese factories, international symbols of mass exploitation, are themselves beneficiaries of cheap goods and consumables, and

⁴⁸ David Jenkins, 'Denying Reciprocity', *European Journal of Political Theory*, 15/3 (2016) pp 312 – 332.

⁴⁹ There are of course a great many people who suffer in ways that cannot even be minimally described as enjoying 'benefits'.

⁵⁰ See Iris Marion Young, *Responsibility for Justice* (Oxford: Oxford University Press, 2011) p 53.

their livelihoods also predicated on the immiseration of others.⁵¹ It is thus possible to simultaneously belong to a group that is exploited, sometimes in appalling ways, and a group that benefits from exploitation: One can both benefit – however minimally – from unjust productive practices, for example, and suffer from them in quite substantial ways.

Individuals are always, as Robert Jubb notes, ‘tragically’ imbricated in processes that sustain injustices, even where the benefits they receive from such involvement are minimal. It is this fact of ongoing and unavoidable involvement in injustice that lies at the root of failures of compliance. Whatever our motivations, we cannot get away from injustice *tout court*. While this of course means we fail to comply with the demands of justice, whether these are conceived at the ideal or non-ideal levels, recognising complicity makes explicit the fact that substantive compliance is always and everywhere impossible. (This notwithstanding the enormous variations in terms of what that involvement actually amounts to, the benefits we enjoy and the burdens we suffer, and how culpable we are for our actions.) However, before attending to some of these variations, we need first to get clear on the nature of this involvement, which I will call complicity.

IV. Complicity

When we discuss the domain of complicity, we are, at the most general level, concerned with the ‘cultural and legal practices, surrounding relations of an agent to a harm that are mediated by other agents’.⁵² We might describe complicity, following Christopher Kutz’s account, as a form of ‘participatory wrongdoing’. We thus need an assessment of what counts as participation in wrong, in order to get clear on whether or not the concept of complicity accurately, and usefully, describes nonideal conditions: We might all be *connected* to severe wrongdoings without thereby *participating* in them. Iris Marion Young’s ‘social connection model’ provides one

⁵¹ Indeed, if you want to know how many slaves you effectively ‘own’, go to <http://slaveryfootprint.org/>. There are 27 million slaves in the world today, the vast majority within – and crucial to – the capitalist system itself.

⁵² Christopher Kutz, *Complicity* (Cambridge: Cambridge University Press, 2000) 2.

means of assessment that will prove useful in determining whether or not our involvement counts as participatory.

Young suggests that ‘individuals bear responsibility for structural injustice because they contribute by their actions to the processes that produce unjust outcomes.’ Participation is wrongful when it results in actions that produce and sustain harms that we can describe as unjust. Reflecting the assumption I make above, ‘our responsibility derives from belonging together with others in a system of interdependent processes of cooperation and competition through which we seek benefits and aim to realize project.’⁵³

This idea of *sustaining* harms, there being a future-oriented component to the relevant actions, complicates matters. Imagine a nurse from The Philippines has the opportunity to come to the UK to work for the NHS. Her family is far from wealthy and taking this opportunity will provide a much needed opportunity to send remittances back home. However, suppose the reason the UK and the Philippines have established this scheme is because of trade links constituted primarily by the sale of military hardware, which have been traditionally used – either today or historically – against the civilian population. Any benefits the nurse gains – whether professional or pecuniary – are predicated on there having been previous outrages, up to and including state-sanctioned murder. However, let’s say that sort of trade has ended and the relationship is now confined to civil purposes. It would no longer be correct to describe the nurse as involved in harm-sustaining actions. She is free from a charge of complicity.⁵⁴

⁵³ Young, *Responsibility for Justice*, 105.

⁵⁴ The same would apply to historic wrongs. However, in so far as these wrongs continue to be part of ongoing injustices – think of capital cultivated out of slavery supplying the wealth and power of many contemporary institutions (Manning Marable, *Living Black History* (New York: Basic Books, 2006) 9 - 12) – then ongoing current actions sustaining injustice are still made possible by not attending to these historic injustices. For any (consistent) libertarian, of course, there should be an even more direct concern with making these wrongful exchanges right, and complicity would seem, from this perspective, to be an accurate reflection of our current situation, since corrupted historic exchanges continue to prevent just exchanges from emerging.

However, on the view here developed, not only would the Chinese workers in the above example be ‘responsible’ for injustices, so too would a slave: A slave, by doing the work demanded of him by his master is also refusing to revolt and upset the structures through which the institutions of slavery continue. He is then, as an implication of Young’s account, ‘responsible’ for the injustices produced by those structures. This might strike us as offensive: How can a person subject to the threat of appalling crimes be held responsible for those crimes? Young in fact accepts this implication, without drawing it to the starkness of this conclusion: ‘An important corollary of the idea that responsibility in relation to structural injustice is shared among all those who contribute to the processes that produce it is that many of those properly thought to be victims of injustice nevertheless share responsibility for it.’ However, though she qualifies her statement by suggesting it is only the ‘many’ who share responsibility, she provides no argument that would show the above starkness, i.e. to include *all*, can be avoided in any principled way. All she can suggest is that any action that would follow from acting on this responsibility involves risks to the bodily integrity of workers and slaves alike, which mean collective action might have to start from somewhere else or await sufficient critical mass.⁵⁵ I believe this is an implication that should be accepted in our thinking through the connections between the possibilities of compliance and complicity, and our understanding of the nonideal/real world that emerges from these connections.⁵⁶

⁵⁵ To quote Joan Robinson, at least with regard to the Chinese worker: ‘The misery of being exploited by capitalists is nothing compared to the misery of not being exploited at all.’ Joan Robinson, *Economic Philosophy* (Oxford: Routledge, 2017 (1962)) 45. Moreover, when these moments of critical mass occur then there does arise a *responsibility* and an imperative to join that mass. This can be stated in some provocative terms: ‘The man who would not fight under our Lord and Master Jesus Christ, in the glorious and heavenly cause of freedom and of God—to be delivered from the most wretched, abject and servile slavery, that ever a people was afflicted with since the foundation of the world, to the present day—ought to be kept with all of his children or family, in slavery, or in chains, to be butchered by his *cruel enemies*.’ David Walker, *Appeal to the Coloured Citizens of the World*, ed. Peter P. Park (University Park: Pennsylvania State University Press, 2000). 14 – 15. At such moments, to *not* follow the imperatives of revolt and resistance is to become morally responsible for the chains by which one is shackled.

⁵⁶ Young, *Responsibility for Justice*, 113.

There are resources within Young's overall account that can help to respond to the concerns raised by the offensive implication of 'victim blaming'. She distinguishes her social connection model from something she calls the 'liability model', an approach she wants to get away from in her discussions of responsibility for injustice. The liability model takes the assigning of blame as the primary goal of assessing responsibility. Rather than focussing on claims to be made against culprits, the social connection model instead looks to the responsibilities people have 'by virtue of their social roles or positions'.⁵⁷ Responsibility is thus given a formal, technical definition in which designations of moral or legal blameworthiness are not the end game. Instead, 'If we contribute by such actions to the processes that produce structural injustice, we are responsible in relation to that injustice. Usually we should not be judged either morally or legally blameworthy or at fault for it.'⁵⁸ There *is* a presupposition that one should not contribute to the sustaining of unjust structures, but there is *not* the additional claim that one is therefore duty-bound to rid oneself of one's complicity, to vacate the sustaining role.⁵⁹

The slave remains included in this idea of 'responsibility', but his 'contributing actions' are shorn of the moral and legal connotations commonly associated with that term: He is responsible in the way a particular cog in a machine is responsible for the functioning of that machine because without it⁶⁰ the machine could not work. There is no reason to suggest this responsibility comes with any notion of blame attached, especially since there are sufficiently good reasons, absent assurances of a collective kind of resistance, to stick to the grind. Any substantively normative claim that might follow on from this distinction, i.e. anything more than a *prima facie* reason to attempt to end the injustices in which one plays a part, must first take into account the locations of the individuals within the process, the resources they have available and the risks they run. The assignment of

⁵⁷ *Ibid* p 104.

⁵⁸ *Ibid* p 107.

⁵⁹ Tamar Shapiro, 'Kantian Rigorism and Mitigating Circumstances', *Ethics* 117/1, 2006. 44.

⁶⁰ Or, at least, without that cogs not-interfering with other cogs in the machine.

responsibility at work in Young's account is more concerned with describing the various components of a particular process, the actions that are required, to keep that process going.

On this more technical view we are, in our roles as both beneficiaries and victims of those processes, unavoidably complicit with processes of injustice. Given the complexity of these interactions, there exists no single agent, even assuming mass coordination, which can take the action needed to prevent injustice and put us, collectively, on a path to justice. To go in the other direction, attempts at refraining from involvement, of making oneself non-complicit or of refusing to benefit from injustice, are impossible. If one takes into account the benefits one receives during immaturity, any attempt to extricate oneself from this tangle of complicity at any one point – perhaps pulling a Han Shan and heading to a mountain monastery – is still not going to scrub a person's hands entirely clean. Those hugely demanding acts of total extrication can only be gestures – of admittedly staggering magnitude.⁶¹ Consequent to all this, demands that emerge from nonideal conditions must take our ongoing and inevitable involvement in injustice as a given, and find ways to *use* this complicity as a means of confronting injustice. It is this which, I argue below, should form the centre of theorising about nonideal conditions, rather than the dubious commentary on motivational paucity that grounds the assumptions of partial compliance.

This understanding of nonideal circumstances offers some improvements on those which take partial compliance as the starting condition. First, it is far more ecumenical about the causes of injustice. It does not pretend to offer a description of exactly *what* we confront in nonideal situations, i.e. the fallout from failures to comply with demands of justice. Instead, it issues only a basic description of one of the key features agents confront in nonideal circumstances, namely, their ongoing and inescapable involvement in processes that inevitably produce injustices. As a consequence of this, it also avoids inadvertently *moralizing* nonideal situations.

⁶¹ A final, practical point: even certain boycotts, where the desire is to refuse consuming goods produced by companies considered especially reprehensible, but which fail to take their cue from the workers who are the intended beneficiaries of such action, can end up hurting the interests of the workers. Irrespective of burdens suffered, the workers still require the livelihoods provided by the injustice-sustaining consumption of their services.

In other words, the imperfections of human agents *qua* human agents are not assumed to be, always, at the causal bottom of our descriptions of nonideal worlds.

However, contrary to Young's intention to move us away from a liability model of responsibility, which she argues is unsuited to the structural nature of injustice, I think that it is possible to use the language of blame to compliment an understanding of responsibility that takes complicity seriously, and goes further, by opening up a way of identifying *culpability*. Indeed, I shall argue below that to do so enables the capturing of some basic features of power and, moreover, to better understand the self-understandings that have motivated, and continue to motivate, those agents and movements responsible for moves to more justice.

A final advantage is that, as a starting point, using this technical idea of complicity to motivate our understanding of the nonideal also points in directions for the means by which we might confront injustice. In short, this is the idea that people's continued and inescapable involvement in the production and reproduction of injustice provides a way of assessing the means by which, cautiously and strategically, injustice might be practically combated.

V: Practice Subversion

For all my scepticism regarding the work partial-compliance can do in describing nonideal condition, I do not wish to entirely abandon discussion of the concept of compliance. Rather, by moving the inevitability of complicity to the centre of our theorising on nonideal conditions, we gain new ways of approaching and understanding the reasons why complying with principles of justice proves so difficult.

In non-ideal circumstances, a person is always, and can only be, non-compliant in the sense that is specified by idealisation, bracketing for now the non-compliance with demands that emerge in the *nonideal* context. This is guaranteed by the impossibility of evading involvement with injustice as just described: One is always going to be performing actions and assuming roles that will produce outcomes in some ways harmful to others.

This kind of complicity is not equally inevitable across all practices within which questions of partial or non-compliance might arise. For an alternative example, consider the issue of compliance with swimming pool regulations. When one enters a swimming pool, one is confronted with a list of the forbidden: bombing, urinating, heavy petting. This kind of compliance is easily accessible because it does not rely on anybody else doing anything in particular – literally everyone else could be breaking those rules and one could still comply with them. Being consistent with the demands of veganism might be another example: one can be a consistent environmentalist, so far as the consumption of meat and meat-related products goes, without relying on similar actions performed

by others (though of course the difficulties/costs of veganism might increase unless sufficient numbers of others take up the cause).

Political, social and economic institutions work in obviously radically different ways from swimming pools: The kind of activity with which one can engage has to make sense both against the background rules constraining certain behaviours, and to what other people are doing within those constraints. The vegan, for example, by paying her taxes might inadvertently be contributing to subsidies for certain kinds of agricultural practices that include animal slaughter. Her otherwise well-intentioned, consistent and high-cost actions are thus undermined, to a degree, by the background against which she makes her decisions. Even she, then, continuing with the more or less technical definition developed above, is complicit in reproducing practices that she considers profoundly unjust, and takes extensive efforts to confront.

Tamar Shapiro's notion of *subversive* non-compliance illuminates the difficulties of practising compliance in non-ideal conditions.⁶² Shapiro imagines a negotiation between two people, one of whom is committed to the practices and intentions associated with negotiation. This includes a willingness to comport oneself in good faith, providing reasons acceptable to the other side and recognising their interests as legitimate. However, the person on the other side of the negotiation lacks such good faith. She treats the negotiations as a means of stalling for time and the other side as an obstacle to get around. Given the others' total lack of interest in negotiation is it still possible to accurately describe the good-faith party as 'negotiating'? There is, Shapiro suggests, a threshold beyond which the others' noncompliance can undermine the basic integrity of the practice such that compliance with it is rendered impossible, whatever the intentions of the properly motivated individual might be.⁶³ The practice is rendered a sham: The attempt to continue on 'to the letter of the law' as defined by the rules of that

⁶² Tamar Shapiro, 'Compliance, Complicity and the Nature of Nonideal conditions', *The Journal of Philosophy* 100/7 (2003) pp. 329 – 255.

⁶³ *Ibid* pp 337 – 338.

practice is to be complicit in that sham. The question becomes what one should do about and *with* one's role inside this sham.

However, there is something in the examples Shapiro uses that tends to obfuscate differences between the practices she considers and the wider issue of complicity with injustice as it confronts most people in the ordinary run of their lives. For instance, a collapse in the negotiations above is a practice from which the parties can pretty costlessly extricate themselves. Other means will have to be found to promote one's ends, but unsatisfactory negotiations can be walked away from. The other example Shapiro provides is of a corrupt police chief, the killing of whom will have, it is assumed, powerful consequences for the internal justice of both the LAPD and, by extension, Los Angeles. Here extrication is a little more difficult – it involves killing someone after all – but again, what is needed as part of the extrication from complicity is a relatively straight-forward and clear cut action, which a single agent can take: Yes, hands will get dirty, but the dirtiness of the murder helps reduce dirt for the practices of the police force overall. In more complex *social* reality, it is impossible to extricate oneself from practices in which adherence to the rules produces injustice and where there is no space beyond those practices into which one can extricate oneself. There is rarely a head that, in rolling, brings us to a substantively better place.

Moreover, the creation of alternative practices or the resistance needed to counter the effects of 'sham' practices, i.e. to get us beyond where we now stand to a more just (set of) practice(s), requires tremendous work. For example, consider the kinds of actions that went into the abolition of the practices surrounding North American slavery. The ending of slavery required, for starters, the emergence of massive internal contradictions between an industrialising north and a still feudal south, as well as individual acts of violent defiance, such as John Brown's raid of the federal armoury at Harper Ferry, triggering a civil war which in its turn led to the deaths of some 620,000. Over and above the de facto ending of slavery, there are the spoils of slavery, which continue to

line the pockets of a great many of America's most venerable institutions.⁶⁴ There is also the violent backlash against the gains of Reconstruction to consider, something that continues to have political, economic and cultural relevance today. Subverting the practices of slavery is incomplete work.

To put this in the context of partial-compliance and complicity, and to bring the point more fully up-to-date, Charles Mills recommends that contemporary white 'race traitors' – whose 'treason is a condition of their loyalty to humanity' – have a choice in speaking out against the consequences of white supremacy, rather than 'unquestioningly "going along with things"'.⁶⁵ However, and Mills recognises this too, unquestioningly or otherwise, "going along with things" – which for much of the time we must – still implies complicity. In other words, challenges to this massively unjust situation are still predicated on accepting a background inability to properly extricate oneself from the sham practices dependent on ongoing injustices.

VI: Constitutive Subversion

Any attempt to collapse this concept of subversion into that of partial or non-compliance is problematic. Compliance, to restate, is not a creative act. It might require creative thought or action in order to complete what is being demanded, but compliance itself represents achieving conformity with given regulations, rules, norms or demands, about which we are aware. Within the practices Shapiro imagines, the prior existence of the practices is taken for granted: In the example of negotiation, both participants *enter into* the practice of negotiation, one of whom behaves in such a way that the practice is undermined. It is thus from within the practice that the sham begins.

⁶⁴See Marable, *Living Black History*, 9 - 12. See also Randall Robinson, *The Debt* (London: Penguin, 2001) p 201 for evidence that a commission that would have been tasked with discovering the *extent* of that lining managed to get just 28 co-sponsors out of a House (of Representatives) of 435.

⁶⁵ Charles W. Mills, *The Racial Contract* (New York: Cornell University Press, 1999) p 107.

This involves an assumption of basic equality between the parties. That is, while there might be differences in power between the negotiating parties, both are considered equal in that both equally possess the quality of having entered into the practice. The good-faith or otherwise is a separate, unequally held characteristic. The power that one side of the negotiating parties might have *vis-à-vis* the regulations and norms prior to entry into the practice is left untheorized. In reality though, those norms and regulations are not simply given as facts by which practices are necessarily constituted but are themselves the consequence of differences in power between differently situated parties. As a consequence of these asymmetries in power, the very constitution of the practice is affected by a kind of ‘shamness’, in which certain agents are able to subvert the norms by which practices are organised *before* questions of compliance or its absence even emerge.

For an instructive example of the subversive constitution of practices, consider different ways of arranging the basic structure of an economy. One such arrangement could be described as corporatist, broadly descriptive of Austrian, Scandinavian, Dutch and German political economies.⁶⁶ Corporatism “involves negotiated bargains among labour, business and government”, which then inform government policy and law.⁶⁷ The arrangement is grounded on a certain level of compromise between these interests, labour promising to refrain from excessive wage demands or strike action in exchange for general levels of state and business support for a redistributivist welfare state and incremental wage increases. This has affects across the organisation of economic resources, working conditions, corporate tax schedules, etc. In such circumstances, and not to overstate the merits of this form of organisation in practice, there are multiple parties representing different interests present at the site of legislative design and thus the constitutive stage of any emergent practices.

Another arrangement is that which John Dryzek labels ‘authoritarian liberalism’. This is a system of governance that excludes other parties – in particular, representatives of labour – from fundamental legislative

⁶⁶In the UK as well, before the 1980’s, a culture of corporatism was captured, colloquially, by the habit of ‘beer and sandwiches at number 10’.

⁶⁷ John S. Dryzek, *Democracy in Capitalist Times* (Oxford: Oxford University Press, 1996) p 43.

decisions, and thus from decisions pertaining to the practices that emerge from them. This alternative arrangement, broadly describable as neo-liberal, reflects a preference – undermined continually in practice – for the market to replace the mode of negotiated compromise that characterises the corporatist arrangement. Indeed, it regards corporatist forms of political economy as interfering with the proper functions of the market, preventing the (necessarily) painful emetic process of flushing away an economy’s inefficient elements.⁶⁸

The corporatist arrangement is not perfect. For instance, it can interfere with more confrontational action that might, at certain junctures, be necessary for change, but which are blocked by labour organisations trying to shore up their place at the table. It might also, conceivably, have certain negative effects on the economy (though in comparison with the “authoritarian liberalist” state, this is increasingly hard to argue).⁶⁹ Nevertheless, what it does allow for is some minimal level of involvement by representatives of groups who are to be affected by the decisions made at that level, and who will have to participate in the practices emanating from those decisions. The equality of entry that Shapiro assumes in her description of practices is thus more closely approximated.

The problem of subversion at the constitutive level is thus, if only partially, remedied by corporatist arrangements: The ‘shamness’ of the practices are mitigated by a more democratic arrangement of the conditions defining those practices in the first place. However, something paradoxical happens at this point. Let us assume the remedy to this constitutive subversion is non-trivial, i.e. under corporatist arrangements the affected parties (labour in this example) have a genuine say in the constitution of the practices with which they will be involved: Conditions are improved, wages are increased, benefits improved, and/or working hours lowered as a consequence of representatives acting with their endorsement. The workers who might previously/counterfactually have suffered from severe exploitation, unstable contracts and exhausting working hours, are now able to enjoy fairer practices with which they can engage in better – if not fully good – faith. Both

⁶⁸ Mark Blyth, *Austerity: A Dangerous Idea* (Oxford: Oxford University Press, 2013) p 81. Of course, the profound *absence* of this emetic process, coupled with the continuing centralising of government power puts the lie to this economic ideology.

⁶⁹ Dryzek, *Democracy in Capitalist Times*, p. 157 (fn 11).

sides – employer and employee – can be said to engage in an arrangement closer to a form of productive *cooperation*.

But such an arrangement will not reduce injustices for such parties as remain outside these constitutive arrangements: The deeper levels of ‘shamness’ remain entirely untouched. Assuming representatives of labour are guided more or less by the material self-interest of their members, then their piece of the pie might get bigger, their conditions might improve, but in other places, perhaps further down the supply chain, injustices remain, injustices often far more substantial than those suffered at this part of the labour arrangements.⁷⁰ All that has happened is that a group previously excluded from decisions regarding the distribution of benefits and burdens, is now complicit in a different, ultimately more culpable way. The inescapability of complicity returns with both renewed force and additional moral baggage.

Any gain of the kind won by labour in the above example, i.e. that precipitates participation in the constitution of certain practices, but which does not take issue with the structures that today still take mass exploitation as foundational, are thereby rendered complicit in the continuation of that exploitation. Moreover, power gained at that constitutive level within such arrangements might also be seen to lead to *culpability*, precisely because one is now privy to decisions that generate the practices that receive as given the wider facts of mass injustice, and upon which those practices depend for their practicability. This leads to the second part of the paradox, this time for parties excluded from the authoritarian liberal arrangement: By continuing to be excluded by the state they avoid complicity at that constitutive level, and thus avoid the culpability this brings about. So long as one is more fully a victim of the constitutive subversion committed by other parties, a person’s complicity cannot develop into that of substantial culpability.

⁷⁰ This is not to neglect the fact that the trade union movement has indeed been a site of solidarity across national boundaries, precisely in recognition of the need for concerted displays of strength against the onslaughts of capital and political power. See Seumas Milne, *The Enemy Within* (London: Verso, 2014) for an account of solidarities within international mining unions.

There are a number of separate questions this essay has pointed toward but which it has not attempted an answer: Is the subversive constitution of practices something for which culpability can ultimately be assigned to one particular agent/set of agents or institution/set of institutions? What is the correct response for people sufficiently moved by subversive practice constitution and their inevitable complicity with them? Can we conceive of even minimal justice until those most excluded from the constitution of practices are given a place at the 'constituting' table? All I have tried to demonstrate is how poor and incomplete the term 'partial-compliance' is as a description of non-ideal conditions. A better description recognises the foundational role played by subversive agency in rendering inevitable the complicity of all with injustice, and the paradoxes and complications associated with the subversive constitution of practices.

It has not been an intention of this essay to advocate for a wholesale *replacement* of the non-ideal question, i.e. what one should do in situations characterised by others not complying with justice. Indeed, it can be conceived as complementary to that project, to distinguish between the various groups of others who are not doing what they should be doing, the reasons why they are not doing it, the benefits accrued by not doing it, and only then to ask the question. The notion of the ultimate culpability of agents engaged in the subversive constitution of practices issues in normative demands that are separate to demands issuing from what one should do when confronted by others' non-compliance.

Conclusion

The use of partial-compliance as an umbrella term for what people and various other agents do in non-ideal circumstances hides a great many distinctions: differences in motivations, differences in complicity, differences in the blameworthiness of that complicity and differences in position within the various hierarchies that constitute an unjust world. With this paper I have tried to show how uncovering these various distinctions leads to a more useful description of non-ideal conditions. While I take no issue with the use of partial-compliance as a general description of the real world, I hope to have demonstrated that absent the distinctions I have here identified, using it as a basis for the normative concerns that are supposed to confront people in non-ideal conditions is of only limited value.

Understanding non-ideal circumstances in a more fine-grained way, complicity is recognised as both inevitable and central. By placing complicity in the wider context provided by the additional consideration of *culpability*, the question as to what a person should do in light of his or her complicity is significantly complicated. So, while every worker and consumer is certainly complicit, and indeed benefits in certain ways from that

complicity, it is important not to let that obscure from view agents who have the power to organise and coordinate, for example, those productive practices described above, or the information to which people have access, and who benefit from it in incomparably greater ways, assuming only trivial costs in so doing. In other words, we should not allow the broad veracity of complicity as a description of people's behaviours and actions prevent us from identifying the agents who are responsible for the *subversive* constitution of practices. Nor should we obscure from view the radical possibility that short of some sort of wholesale revolution all that is possible is widening the range of culpable agents, while ensuring that substantive justice remains a dim prospect.

5. Weaponising Justice and Equality

WORKING PAPER by Hwa Kim Young (University of Warwick)

Introduction

How should we convert the principles of justice developed by political theorists into practice? The most explicit path is set out by John Rawls. We start off with an ideal theory that sets out the ‘ultimate target’ that we should aim for (Stemplowska and Swift, 2017). It does so by aiming “to identify a set of stringent principles grounded on a plurality of values and set of factual assumptions that can guide us in designing and maintaining our most profoundly influential social institutions” (Williams, 1998: 124). Only after having developed this ideal, do we move on to non-ideal theory. Non-ideal theory provides more action-guiding principles that act as a roadmap towards the target provided by ideal theory, telling us what to do to get there. Normative theory, at the non-ideal stage, is more contextual and empirically-sensitive to real-world conditions, and acknowledges that not everyone will act justly. For Rawls, then, one central task of non-ideal theory is to answer the question “what shall we do here and now given the ideal endpoint that we want to reach?” (Rawls, 1999a: 210).

I believe there is something missing from this model of theory and practice. Unlike in mere theoretical discussion, political struggle is arduous and involves significant stakes. Political agents need to compete with unreasonable people who are willing to use unreasonable methods to fight for their ideals in the arena of politics. A set of ideal principles will have little social and political impact if most political agents do not feel compelled to fight for them in this context. A set of non-ideal action-guiding principles that act as a roadmap to the ideal will have little impact if no one feels compelled to read the map or act in the way that the more action-guiding

principles want us to. Both ideal and non-ideal theories must have components that inspire people to fight for justice in the face of difficult political opposition.

G. A. Cohen provides another account of the relationship between theory and practice that provides a solution to this problem. Rather than providing us with a systematic account of our considered judgments of the just and unjust and how this should guide our design of our most profoundly influential social institutions, Cohen theorises about justice in a way that ascertains “the very nature of distributive justice” in its purest form (Cohen, 2008: 301). Alan Hamlin and Zofia Stemplowska refer to this manner of theorizing as the ‘theory of ideals’, which aims to “identify, elucidate, and clarify the nature of an ideal or ideals” (Hamlin and Stemplowska, 2012: 53).⁷¹ Cohen’s exclusion of feasibility, publicity, and stability from justice suggests that he would be the last person to turn to when thinking about justice and its impact on social and political practice. However, one of the reasons why Cohen excludes these features from justice is that he holds a different model of theory and practice. Theory is not meant to “generate a comprehensive social design which the politician then seeks to implement” (Cohen, 2011: 212). It should not aspire to do this because politics is an endless struggle, which makes the full achievement of justice impossible. Instead, for Cohen, theory “serves as a weapon” in this endless struggle (Cohen, 2011: 212). Justice should provide the “depth of conviction” necessary for political agents to push for social and political change in the real world, which can only come from “*theory that is too fundamental to be practicable in a direct sense*” (Cohen, 2011: 213). I will refer to this as the conviction model of theory and practice.⁷²

How we theorise about justice can have knock on effects on how justice can generate political impact. The difference between these models suggests there is a tension in the way we can think about the relationship between theory and practice. Either we have a theory of justice that provides us with a systematic basis of our

⁷¹ They also argue that there is “simply no reason for the theory of ideals to take account of issues of feasibility, since the inquiry is into the nature and structure of the normative criteria to be employed”, and not about institutional design (Hamlin and Stemplowska, 2012: 55).

⁷² I will also refer to it as the conviction role of justice or ‘weaponising’ justice

considered judgments which we can use to make principled trade-offs, or we have a hyper-idealistic theory of justice that can generate the depth of conviction necessary to inspire political agents to fight for justice in arduous political struggle; in other words, a tension between the ‘systematicity’ of a theory and the political compellingness of a theory. This tension is not limited Rawlsian political philosophy; it applies to anyone who wants to provide (a) an ideal theory that provides “the only basis for the systematic grasp of [the] more pressing problems [of non-ideal theory]” and/or (b) a non-ideal theory that is more ‘realistic’ than ideal theories, in the sense that it is more contextual to the real world here and now, and is more attuned to social-scientific data (Rawls, 1999a: 8). Rawls acts as the best representative of these two positions, but my analysis also applies to some of the work done by non-ideal theorists such as Elizabeth Anderson and Amartya Sen.⁷³ Is there a way political philosophers can generate depth of conviction while also providing a systematic basis of our considered judgments?

The first section of this paper will provide a more comprehensive account of the conviction model of theory and practice. I will then discuss the limitations that Rawlsian constructivists face when it comes to weaponising justice, namely the complexity of its moral methodology, its concern with feasibility, and its inherent desire to establish consensus. These limitations are meant to show that Rawlsian ideal theory alone cannot play the conviction role as effectively as Cohen’s approach to justice can. I will then discuss a major Rawlsian response, namely that some political agents might be more inspired to fight in political struggle if they knew their conception of justice was feasible (in the long run). I show how this response is sound, but it also applies to Rawls’s theory of justice and its ability to provide depth of conviction for political agents. I will argue that

⁷³ Sen advocates a comparative approach among the different options within a feasible set of policies, which requires a lot of empirical evidence about the opportunities available to us here and now (Sen, 2009). Anderson argues that her work “integrates research in the social sciences in ways not ordinarily found in works of political philosophy” (Anderson, 2010: 3). In this paper, I focus on Rawls versus Cohen because (a) “Much academic work in political theory today ... operates in a Rawlsian paradigm” (Swift and White, 2008: 59) (b) a lot of Cohen’s meta-ethical arguments against Rawls about justice (and its ‘fundamentality’) can easily be converted to arguments about political impact and (c) it is surprising to interpret Cohen’s work as being in one important way more politically realistic than Rawls’s realistically utopian theory of justice.

Rawlsian ideal theories, which provides the systematic theory to balance competing values, should (ideally) be accompanied by a Cohenite ‘theory of ideals’ that can help political agents effectively deliberate and engage in political struggle. Combining these two approaches can reconcile the tension between the ‘systematicity’ of a theory and the ability for that theory to generate the depth of conviction necessary to inspire political agents in political struggle.

The Conviction Model of Theory and Practice

In the methodological debates in political philosophy, Cohen is most well-known for his attempt to ‘rescue’ justice and equality from Rawls’s constructivism and its inclusion of facts and constraints of publicity and stability. His rescue of justice rejects the Rawlsian belief that justice should be ‘realistically utopian’. He explicitly states “it is worth pointing out that [Rawlsians] are in one way more Utopian than I. For in believing that justice must be so crafted as to be bottom-line feasible, they believe that it is possible to achieve justice, and I am not so sanguine” (Cohen, 2008: 254). Some might argue we should not turn to a theorist who expresses their belief that justice is unachievable for advice on how to convert theory into practice.

This would be true if we neglected Cohen’s alternative model for the relationship between theory and practice. We can go even further by arguing that one of the main reasons why Cohen’s approach to justice is intuitively appealing (relative to Rawls’s) is the model of theory and practice it holds. Feasibility is unimportant for theory in this model because it should not be trying to generate a comprehensive social design. He argues that “things don’t work that way, because implementing a design requires whole cloth, and nothing in contemporary politics is made out of whole cloth” (Cohen, 2011, 212). According to Cohen, politics is an endless struggle, which makes the full implementation (and maintenance) of any theory of justice impossible. Political agents engaged in political struggle do not have the luxury of only deliberating with reasonable people; they must compete against

those who are unreasonable, both in the sense that they hold incorrect normative and empirical views, and are willing to use unreasonable methods to advocate their political positions.⁷⁴

Instead, theory should instead act as a weapon, in the endless struggle of politics because “it provides a characterisation of its direction, and of its controlling purpose” (Cohen, 2011: 212). To serve this function, theory needs to be “uncompromisingly fundamental ... not devised with one eye on electoral possibility” (Cohen, 2011: 212). Theory expresses ultimate convictions which inspires political agents to continue fighting in the endless struggle of politics, even though they know they will never fully succeed. Cohen uses the example of “the theories of Friedman, Hayek, and Nozick” (Cohen, 2011: 212). If we consider their theories to be “practical proposals”, they were “crazy in the strict sense that you would have to be crazy to think that such proposals (e.g., abolition of *all* regulation of professional standards and of safety at work, abolition of state money, abolition of *all* welfare provision) might be implemented in the near, medium, or long term” (Cohen, 2011: 212). Instead of practical proposals, we should think of these theories as inspiring political agents to fight for their positions in politics. As Cohen argues “politicians and activists can press not so crazy right wing proposals with conviction because they have the strength of conviction that depends upon depth of conviction, and depth comes from theory that is too fundamental to be practicable in a direct sense” (Cohen, 2011: 212-213).

For Cohen, fundamental theories provide principles that are ‘robust’ and ‘pure’. Williams describes fact-insensitive principles as more robust compared to fact-sensitive principles because the justification of fact-sensitive principles “always depends on the truth of contingent empirical assumptions” (Williams, 2008: 115). Principles of justice are ‘pure’ in the sense that they “answer only to one distinctive value, namely justice” and do not have to answer to ‘alien factors’ such as stability, publicity, and feasibility (Williams, 2008: 115-116). These features are important for theories that act as weapons in the struggle for politics, as they represent “the summit of our normative convictions”, which we will rely on to continue motivating us in endless and rigorous political

⁷⁴ In the methodological debates in political philosophy, Cohen is usually considered to be even more idealistic than Rawls. However, when it comes to the domain of politics, Cohen seems to have a similar account offered by political realists. See (Rossi and Sleat, 2014).

struggle (Cohen, 2008: 20). Furthermore, for Cohen, the justification of less robust and ‘impure’ principles ultimately depend on appeal to robust and pure principles.⁷⁵

Theory that is too fundamental to be practical in a direct sense also provides political agents with “justificatory rhetoric ... to push for (or resist) change” (Cohen, 2011: 213). Cohen argues that “all change in modern conditions ... is perforce incremental, 2 percent here, 5 percent there, accumulating after, say, fifteen years, into a revolution” (Cohen, 2011: 213). Fundamental theory provides the conviction along with the resources to produce powerful justificatory rhetoric to attack opponents, generate solidarity amongst supporters, and generate support from those who are undecided. Truly fundamental theory (and the justificatory rhetoric it produces) is “needed to defend every half-mile of territory gained and to mount an attempt to regain each bit that has been lost” (Cohen, 2011: 213).

Finally, depth of conviction requires affirmation of traditional partisan principles. Cohen argues that Hayek, and Friedman did not offer fundamentally new ideas. “Instead, they explored, developed, and forthrightly reaffirmed the Right’s traditional principles” (Cohen, 2011: 211). This provides an ideological self-confidence necessary for political success. According to Cohen, the Labour Party historically “affirmed a principle of community and a principle of equality ... They were, indeed, the only values which *the Left affirmed as a matter of principle and which the Centre and Right reject as a matter of principle*” (Cohen, 2011: 214). Reaffirming traditional partisan principles serves at least two purposes. The first is that it is the best way to prevent you from implicitly endorsing the standards of your political opponents, which Cohen believes the Labour party did with the tax-restraint of the Conservatives. Secondly, ideological self-confidence provides a clear basis for voters to vote for your political party. If the major political parties all accepted the same fundamental values, the only basis in which voters could choose who to vote for would be technocratic competency. Not only might this lead to voter

⁷⁵ Cohen argues that “all principles that reflect facts reflect facts only because they also reflect principles that don’t reflect facts, and that the latter principles form the ultimate foundation of all principles, fact-reflecting principles included” (Cohen, 2008: 254).

apathy but without ideological self-confidence, there is little reason to believe that Left-leaning parties will in fact be more technocratically competent than their ideologically self-confident opponents.

I think we can find this model of theory and practice throughout Cohen's work. It allows him to write a short and concise book called 'Why Not Socialism?' in which he portrays a beautiful picture of life in a camping trip arranged by socialist principles and where people are motivated by values of community and solidarity, without having to worry about the independent question of feasibility.⁷⁶ I think it is relatively uncontroversial to argue "that all people of goodwill would welcome the news" that everyone treated everyone else with "the reciprocating attitude that is characteristic of friendship ... it is surely a welcome thing when more rather than less community is present in society" (Cohen, 2009: 51-52). When criticising Rawls, he argues that when he says "'x represents an injustice' ... It will not mean 'x represents an injustice that ought to be rectified by the state' ... It will mean, more elementarily, that the world is less than fully just by virtue of the presence of x in it" (Cohen, 2011: 83). Cohen is (at least partially) motivated to provide a beautiful picture of how the world should be, without this source of inspiration being burdened by distinct considerations of feasibility, stability, or what our fundamental social institutions should look like when they structure the distribution of benefits and burdens. Finally, he praises Dworkin's luck egalitarianism by pointing out its impact on weaponising egalitarian justice. Cohen famously claimed, "Dworkin has, in effect, performed for egalitarianism the considerable service of incorporating within it the most powerful idea in the arsenal of the antiegalitarian Right: the idea of choice and responsibility" (Cohen, 2011: 32).⁷⁷

⁷⁶ He does have a section in the book that asks whether the socialist ideal is feasible. However, he doesn't directly answer the question (nor did he intend to).

⁷⁷ Someone might argue that Cohen's attempt to rescue justice and equality is purely intellectual and his model of theory and practice is irrelevant to this intellectual project. I don't think this is accurate. He explicitly states in his book *Rescuing Justice and Equality* that he is an 'Oxford man' and Rawls is a 'Harvard' man, which was his shorthand for stating that "we have different views about the analytic/synthetic distinction, and therefore about the status of philosophy itself; also about fact and value, and about the relationship of theory to practice" (Cohen, 2008: 3-4). However, my aim is not to provide the most accurate portrayal of Cohen's motivations. My aim is to show how Cohen's intellectual endeavours can be made more

(footnote continued)

To summarise, for Cohen, theory should provide egalitarians with a depth of conviction, which can only emerge from affirming uncompromisingly egalitarian fundamental principles. Theory that does this can be used as a weapon for political agents to (a) inspire political agents to continue fighting in the endless struggle of politics, (b) generate justificatory rhetoric to attack political opponents and attract supporters, and (c) affirm traditional egalitarian principles that provide ideological self-confidence.⁷⁸ I refer to this as the conviction model of theory and practice.⁷⁹

Can Rawlsian theories of justice fulfil this conviction role? The quick Cohenite answer is no. Clearly Cohen believed that what Rawls was discussing were not fundamental principles of justice, but ‘rules of regulation’.⁸⁰ However, it would be unfair to argue that Rawls constructed his principles of justice with ‘one eye on electoral possibility’. Whether Rawls’s principles should be considered ‘fundamental’ or can somehow be converted into fundamental principles is a controversial issue (Pogge, 2008). Even Cohen recognises that the distinction between fundamental normative principles and rules of regulation “is not exhaustive because there exist derivative normative principles, some fact-insensitive and some not, which are not rules of regulation” (Cohen, 2008: 276). This suggests that Rawls’s principles of justice might fall between these extremes. The quick answer seems a little unfair to Rawls’s theory; I believe that it can play the conviction role. However, there are some limitations that Rawls’s theory faces due to the constructivist moral methodology he uses.

attractive to someone who is concerned with political practicability (who might immediately dismiss Cohen as being far too idealistic), and show what impact this model of theory and practice has on Rawlsian justice.

⁷⁸ It is important to note that political agents and politicians can have ‘strength’ of conviction for policy proposals and rules of regulation. However, the strength of conviction is warranted only if the policy proposals and rules of regulation are properly justified by more fundamental principles of justice. The ‘depth’ of conviction can only come from fundamental theory.

⁷⁹ I also use the terms ‘conviction role’ and ‘weaponising’ justice throughout the paper.

⁸⁰ According to Cohen, fundamental normative principles “are not derived from *other* normative principles”, whilst rules of regulation are rules that “we *adopt* ... to order our affairs: we adopt them in the light of what we expect the effect of adopting them to be”. (Cohen, 2008: 276)

Limitations to Weaponising Rawlsian Justice

What are the limitations of weaponising Rawlsian justice? In this section, I raise three of the main limitations. The first is the complexity of the moral methodology set out by Rawlsian constructivism. This complexity plays a major role in limiting the justificatory rhetoric that theory should provide. The second limitation is the inclusion of feasibility constraints into justice. Cohen believes that this inevitably ‘waters-down’ the depth of conviction for any theory of justice. Finally, I discuss how the constructivist moral methodology is inherently focused on consensus building, which makes it difficult to weaponise.

Can we write a ‘Why Not Rawlsianism?’

The complexity of Rawls’s enterprise undermines its effectiveness in political messaging and deliberation. The problem of complexity can be seen in both the problem that Rawls is trying to solve and his method. For Rawls, his aim is to define “the fundamental terms of ... association” and to “specify the kinds of social cooperation that can be entered into and the forms of government that can be established” (Rawls, 1999a: 10). This is an incredibly complex problem that “is bound to affect the problems of efficiency, coordination, and stability” (Rawls, 1999a: 6). We need to know what distributive role justice is playing, but also what kind of broader consequences that justice might have. Cohen is interested in “the very nature of distributive justice” (Cohen, 2008: 301). However, he also recognises that the Rawlsian question is much more complicated and requires many considerations other than justice. He argues “justice, whatever its content may be, is only one thing to be taken into account in devising sound rules of social regulation” (Cohen, 2008: 302). To put it simplistically, someone making an all-things-considered judgment will have a more difficult time than someone analysing one value.⁸¹

⁸¹ Cohen himself faces this issue when he needs to trade-off justice with other values. I will discuss this further in my section on Reconciliation.

The complexity of the method emerges from the fact that there are multiple steps in the construction process of justice that all need to fit together into a coherent package. For Rawls, what is distinctive about his moral methodology is that “it specifies a particular conception of the person as an element in a reasonable procedure of construction, the outcome of which determines the content of the first principles of justice” (Rawls, 1999b: 304). Rawls conceives persons as “rationally autonomous agents of construction” who “view themselves and one another in their political and social relations as, free and equal moral persons” (Rawls, 1999b: 308-309). The fair procedure that represents these citizens as free and equal is what he refers to as the ‘Original Position’. Rawls states that one of the important features of the original position is that “no one knows his place in society ... This ensures that no one is advantaged or disadvantaged in the choice of principles by the outcome of natural chance or the contingency of social circumstances” (Rawls, 1999a: 11). These are just the initial steps of the constructivist moral methodology. I think the complexity of the method is warranted from the perspective of moral justification. It is required to provide a “systematic account of our considered judgments of the just and unjust” (Rawls, 1999a: 36). This can be a major source of conviction for those who are willing and able to take the time to carefully study and assess Rawls’s work.⁸² However, in political struggle, we do not have the luxury of only engaging with reasonable people who are willing and able to walk down the deliberative path set out by Rawlsian constructivism. There are many who hold mistaken empirical and normative views and are unreasonable both in Rawlsian and conventional ways. The way that political struggle occurs in our media today (and for the foreseeable future) suggests that there is little time and demand for long-winded moral justifications, even if they happen to be sound. What this suggests is that Rawlsian justice can only provide depth of conviction for a subset of people, which is insufficient for generating substantial political change.

The problem of complexity is not just limited to the idea of whether ‘regular’ political agents will engage in the process of moral deliberation set out by Rawls. The complexity makes it very difficult for the development of justificatory rhetoric that theory should provide. The proper relationship between theory and justificatory

⁸² I will discuss this idea further in the section on Reconciliation.

rhetoric can be seen in Cohen's work and his *Kidnapper's Argument* against incentives. Cohen describes a case in which a kidnapper captures a child and is demanding ransom money from the parents of the child. The kidnapper tells the parents "children should be with their parents, unless you pay me, I shall not return your child. So you should pay me" (Cohen, 2008: 39). He states that the minor premise of the argument 'unless you pay me, I shall not return your child' is made true by the kidnapper, which is one of the reasons why we find this statement made by the kidnapper despicable. To put it simplistically, Cohen argues that the rich and talented are doing something similar when they demand extra payments to incentivise them to work harder in a way that will benefit the worst off in society. He imagines a set of rich managers addressing poor workers. The managers argue "public policy should make the worst off people (in this case, as it happens, you) better off. If the top tax goes up to 60 percent, we shall work less hard, and, as a result, the position of the poor (your position) will be worse. So the top tax on our income should not be raised to 60 percent" (Cohen, 2008: 59). What this argument suggests is that the rich managers are unwilling to do what they could do to make the worst-off better off, even when the managers would be much better off than most. Cohen argues that the rich managers cannot claim that "the money inequality that they defend is necessary to make the poor better off, since it is they who make it necessary" (Cohen, 2008: 60). This suggests that the "incentives argument can justify inequality only in a society" where "there is a significant lack of community", which is true in the kidnapper case (Cohen, 2008: 47). My aim is not to argue that Cohen has the most plausible account of the relationship between justice and incentives. I raise this issue because Cohen is right to state that in today's political discourse, "real world inequality ... is actually defended on incentive grounds" (Cohen, 2008: 56). Cohen's fundamental justice allows him to generate powerful justificatory rhetoric (rich people demanding incentive generating levels of tax are acting like kidnappers) against those who consistently argue that it is in the interest of everyone to lower taxes for the rich, which is an argument that is almost accepted as a fact in regular political discourse today.⁸³

⁸³ On the level of public policy, Cohen and Rawls might reach the same conclusion. Cohen states that "it might [nonetheless] be bad policy to level down" (Cohen, 2008: 319). However, Cohen's theory has more resources to generate justificatory rhetoric against these types of arguments, whilst Rawlsians generally accept the idea that deviations from equality can be justified for certain reasons.

Can we generate powerful justificatory rhetoric for Rawlsian theories of justice? I believe we can. However, unlike Cohen's theory of justice, there is a trade-off that Rawlsian constructivists face. Rawlsians can either generate a rhetorically attractive argument that inaccurately portrays the theory, or it can generate an accurate account of the theory without the justificatory rhetoric. An example will help clarify this trade-off. The original position is an extremely influential thought experiment. This might be a good place to start for Rawlsian justificatory rhetoric.

When it comes to justificatory rhetoric from the Original Position, it might be an effective tool in the struggle of politics to ask people who disagree with basic egalitarian commitments the question 'imagine if you were in the original position ...' or 'would you still believe that if you were in the original position?' However, at best, this seems like an invitation for further thought and engagement with the best interpretation of the public political culture of liberal democratic society. It doesn't seem to have the same sharp edge that 'you are acting like a kidnapper' does. Considering how the original position is used "to help us work out what we now think", this limitation is not surprising (Rawls, 1999b: 402). A more forceful argument derived from the Original Position might be that 'morally arbitrary features should not affect the fair distribution of benefits and burdens in society'. This might be an argument that one can advance from the perspective of the Original Position that is more substantial than inviting the opponent to think again. However, the problem with this version of justificatory rhetoric is that it provides an inaccurate account of Rawls's comprehensive theory of justice.⁸⁴ This statement seems to be more representative of luck egalitarianism than it does Rawls's two principles of justice.

⁸⁴ Someone might respond by stating that the justificatory rhetoric that emerges from the incentives argument (you're acting like a kidnapper when demanding lower taxes) also does not accurately portray Cohen's theory because it is only one argument. This is literally true in the sense that Cohen's theory of justice is not just about incentives. However, the incentives argument only works as part of a theory of justice that is radically egalitarian. A libertarian like Robert Nozick would not be able to make the incentives argument. Even though the incentives argument is just one argument made by Cohen, it still provides an accurate vision of what Cohen's theory of justice demands, namely radical egalitarianism. The same is not true for the examples I use from Rawls's theory.

The concept of a well-ordered society might be a good way of providing a beautiful picture in a similar way that Cohen does in ‘Why Not Socialism’, whilst at the same time being comprehensive in a way that provides an accurate picture of Rawls’s theory. Rawls defines a well-ordered society as “one designed to advance the good of its members and effectively regulated by a public conception of justice ... it is a society in which everyone accepts and knows that the others accept the same principles of justice, and the basic social institutions satisfy and are known to satisfy these principles” (Rawls, 1999a: 397). This provides a partial account of the end of social unity that Rawls wants his principles of justice to achieve. There is something appealing about this picture. There is a sense in which we jointly pursue justice as a shared “very basic political end” (Rawls, 1996: 202). The fact that everyone accepts (and knows that everyone else accepts) the same principles of justice suggest that there is a sense of civic friendship that Cohen would also appreciate. However, this partial picture requires a lot more details filled out by what the principles of justice are. Once we provide these details, and the moral justifications behind them, we reach a comprehensive account of Rawls’s theory. However, I don’t think this can play the desired justificatory rhetorical role in political struggle (at least not as effectively as the kidnapper argument from Cohen’s theory).

The important point is not that these examples are the best forms of justificatory rhetoric that can be derived from Rawls’s theory of justice. There will most likely be a political rhetorician or strategist who can do a better job than I have. The point is that they will face a trade-off. Either it will provide a sharp and direct weapon in the political arena at the cost of providing an accurate picture of the theory, or it will provide a comprehensive account of the theory which waters down the directness of the justificatory rhetoric.

Feasibility and Watering Down the Depth of Conviction

Cohen has consistently argued that Rawls’s inclusion of conditions such as feasibility, publicity, and stability in the construction process of justice suggests that what Rawls is constructing are not fundamental principles of justice. He finds the question “what ought justice, or the principles of justice, to be?” and speaking of concepts of “‘constraints on’ and ‘desiderata of’ justice to be incoherent” (Cohen, 2008: 277). Cohen made these arguments against the meta-ethical status of Rawls’s principles. However, we can apply them to our capacity to weaponise Rawlsian justice. Cohen uses an interpretation of Rawls provided by Murphy and Nagel to question Rawls’s commitment to his commitment to his two principles of justice, particularly the difference principle. Murphy and Nagel argue “[Rawls’s] most famous and controversial claim is that differences in natural ability—the inequalities of what he calls the natural lottery—have a morally arbitrary effect when they result in differences of earning power. He holds that since no one can be said to deserve the genetic endowment they are born with, desert

cannot justify that proportion of difference in material rewards between an unskilled labourer and a highly trained professional that might be attributable to the difference between their genetic endowments. This leaves only people's free choices as a possibly nonarbitrary source of inequality, and Rawls is sceptical that there are feasible institutions that could detect the extent to which people's fortunes are due entirely to their choices. That in turn leads him to support the difference principle, which requires the elimination of all inequalities up to the point where greater equality could only be achieved at the cost of harming the worst off" (Murphy and Nagel, 2002: 55).

I do not think that the main reason Rawls has for supporting the difference principle is pragmatism. However, it is easy to see how someone like Cohen could use this interpretation to argue that "If Murphy and Nagel are right, then it might be said that Rawls himself identifies justice with the luck-egalitarian principle, but that he supports the difference principle because justice itself (so conceived) is infeasible" (Cohen, 2008: 271). What this suggests is that the depth of conviction lies not with the difference principle as a principle of justice, but with the luck egalitarian principle. Luck egalitarianism represents the 'summit of Rawls's ultimate conviction'. Rawls's depth of conviction lies in the luck egalitarian principle, which he uses to justify the difference principle, which means that the strength of conviction that Rawls has for the difference principle is the outcome of the depth of conviction provided by the more fundamental luck egalitarian principle of justice. Rawls's advocacy of the difference principle represents a watering down of his ultimate conviction, due to concerns of feasibility. For Cohen, depth of conviction for principles will inevitably be watered down for anyone who imposes feasibility constraints on their principles of justice.

Weaponising Consensus?

The desire for consensus in Rawls's theory is most evident in his later work and his desire for an overlapping consensus as the basis for a well-ordered society. However, I am more concerned with the consensus-building inherent in the moral methodology of constructivism and what impact this has on weaponising justice.

In what way is constructivism as a moral methodology reliant on consensus? The very idea of Rawlsian justification relies on it. As Rawls argues, the "social role of a conception of justice is to enable all members of society to make mutually acceptable to one another their shared institutions and basic arrangements, by citing what are publicly recognised as sufficient reasons, as identified by that conception. To succeed in doing this, a conception must specify admissible social institutions and their possible arrangements into one system, so that they can be justified to all citizens ... the task of justifying a conception of justice becomes: how can people settle

on a conception of justice, to serve this social role, that is (most) reasonable for them in virtue of how they conceive of their persons and construe the general features of social cooperation among persons so regarded?” (Rawls, 1999b: 305). The very idea of objectivity in Rawlsian constructivism is based on “the search for reasonable grounds for reaching agreement rooted in our conception of ourselves and in our relation to society ... what justifies a conception of justice is not its being true to an order antecedent to and given to us, but its congruence with our deeper understanding of ourselves and our aspirations, and our realisation that, given our history and the traditions embedded in our public life, it is the most reasonable doctrine for us” (Rawls, 1999b: 306-307).

There are many more aspects of Rawlsian constructivism that suggests it is inherently consensus-building. The multiple formal constraints of the concept of right can also be considered integral for this purpose. However, the issue is not how many different elements of Rawlsian constructivism is designed with consensus-building in mind. The issue is what impact this has on the ability for Rawls’s theory to act as a weapon against political opponents in political struggle. I believe there are at least two implications. Firstly, there is something hypocritical about using a consensus-building mechanism as a weapon against others. This would especially be the case against a political opponent who initially disagreed with the Rawlsian position, but eventually thought about changing their mind. The more important implication which Cohen might have in mind is the negative incremental impact it has on political struggle. To understand this implication, we need to remember how Cohen believes justice improvements occur. As mentioned before, Cohen believes that politics is an endless struggle. Political and social change occurs in an incremental basis; “2 percent here, 5 percent there, accumulating after, say, fifteen years, into a revolution” (Cohen, 2011: 213). At the same time, there is no guarantee that there will be an incremental improvement in any year, and political agents will need to constantly resist forces that attempt to generate incremental change in the opposite direction. This is one of the most significant ways in which justice acts as a weapon, “it provides a characterisation of its direction, and of its controlling purpose” (Cohen, 2011: 212).

A Rawlsian constructivist and their desire for principles that can achieve a consensus will inevitably include different features into their construction process. Reasonable people with different perspectives might accept that equality is an important value, but might find it less attractive if we didn’t introduce features such as pareto-optimality or efficiency into our conception of justice. Cohen believes that alien factors can water down the fight for traditional left-wing principles of equality and community. When it comes to the best/most reasonable conception of justice, Rawlsians might be right to include these alien factors into their moral deliberation. However, from the perspective of political struggle, it can undermine the ideological self-confidence of the Left. The Right is also fighting for the values of efficiency and responsibility. This means there is a larger

coalition of people fighting for these factors, providing a greater force for incremental change in the Right's preferred direction. Egalitarians have good reasons not to contribute to this incremental change in the direction of their political enemies. A moral methodology that is inherently consensus-building might have the consequence of unintentionally contributing to this incremental process in a negative manner.

Does feasibility necessarily undermine desirability?

How Cohen engages in political philosophy is much more intuitively appealing once we understand the model of theory and practice he has in mind. However, there is a major response that Rawls can provide against the Cohenite intuition that Rawls's theory cannot provide the depth of conviction necessary for political agents. For some political agents, the motivation to fight for justice will only emerge if they know that justice is (in the long run) feasible. I believe this response is sound. However, the response is limited in its ability to criticise Cohen's theory of justice, as it can also apply to Rawls's theory. I will start by describing the response before moving on to how it also applies to Rawls's theory.

Not everyone will believe that concerns of feasibility water down the desirability of a theory of justice. The fact that there has been an explosion in the literature on the methodological disputes in political philosophy suggests that there is a significant demand for more 'realistic' normative theories. As David Miller argues "If justice cannot be achieved on earth, and there is no afterlife to look forward to, then there is nothing left for political philosophy *but* lamentation over the size of the gap that unavoidably exists between the ideals it defends and the actual conditions of human life" (Miller, 2013: 231). I believe some political agents will want something more from political theory than "to discover what we have reason to regret" (Miller, 2013: 231). For people who hold this view, a theory of justice that proved it was both feasible and desirable would provide the depth of conviction that Cohen believes is so important for political agents to continue engaging in the difficult political struggle. In fact, Cohen's view about justice (as one value of distributive justice, and not as an all things considered view of our entitlements) allows for the "conceptual possibility of an unjust society in which nobody is doing anything wrong ... To think of a society as unjust when, for example, individuals are not getting their fair distributive share, and the reason they are not getting it is that we cannot make it happen that they do ... is to create a gap between the idea of injustice, on the one hand, and the idea that some are treating others unjustly, on the other" and I think Stemplowska and Swift are right to argue that "some will find that too high a price to pay" (Stemplowska and Swift, 2017: 386).

There are people who do not share Cohen's view that feasibility 'waters down' the overall desirability of a theory of justice. Rawls's justice is better placed to generate the depth of conviction necessary for these political agents to fight in political struggle. However, this Rawlsian response also applies to Rawls's theory of justice. Even for people who believe that justice should ultimately be feasible, Rawlsian constructivists face two limitations for providing depth of conviction. The first limitation emerges by questioning how realistic Rawlsian 'realistic utopias' are. The second can be seen by questioning what we mean when we say theories have feasibility constraints.

Is Rawls's conception of a realistic utopia actually 'realistic'? It is important to note that realistic utopias in the way that Rawls defines them are not necessarily realistic in the sense that "its realisation ... [is] empirically likely, even in the very long term" (Stemplowska and Swift, 2017: 380). There are several conditions that a conception of justice needs to meet for it to be considered realistically utopian. Generally, Rawls emphasises stability. From *A Theory of Justice*, Rawls argues that parties in the original position should choose principles that "generate its own support. Its principles should be such that when they are embodied in the basic structure of society men tend to acquire the corresponding sense of justice and develop a desire to act in accordance with its principles" (Rawls, 1999a: 119). The issue is not 'how do we reach a society governed by my two principles of justice?' The issue seems to be 'if the basic structure were to be regulated by my two principles of justice, would citizens with a sense of justice abide by those principles over their own self-interest?' Rawls's discussion in *Political Liberalism* provides a similar account of a realistic utopia.⁸⁵ When asking the question 'How is Political Liberalism possible?' Rawls asks the question "how can the values of the special domain of the political – the values of a subdomain of the realm of all values – normally outweigh whatever values may conflict with them? Put another way, how can we affirm our comprehensive doctrine and yet hold that it would not be reasonable to

⁸⁵ Cohen focused his attention on *A Theory of Justice* rather than *Political Liberalism*. He states that he "for the most part ignore late Rawlsian theses about the 'politicality' of justice, not because I am out of sympathy with them (although I am), but because my critique of an essentially *Theory*-centered conception of justice is pretty complex as it stands, and it seems wise to set aside the special task of measuring how well that critique also applies to late Rawlsian theses" (Cohen, 2008: 298). However, I discuss political liberalism here because (a) it provides a clearer account of what Rawls meant by 'realistic utopia' and (b) this paper is not directly engaged in the Cohen versus Rawls debate on metaethics.

use state power to gain everyone's allegiance to it?" (Rawls, 1996: 139)⁸⁶ Rawls refers to this type of stability as 'stability for the right reasons'. Once this possibility is shown, Rawls believes that this is sufficient for his conception of justice to play the role that "Kant gave to philosophy generally: the defense of reasonable faith ... in our case ... reasonable faith in the possibility of a just constitutional regime" (Rawls, 1996: 172). Like Rawls, I believe this is sufficient for ideal theory. However, for those who believe that feasibility increases the depth of conviction they have for a conception of justice, some may find that reasonable faith of this kind may be insufficient. They might argue that for a feasible theory to provide the depth of conviction necessary to motivate political agents to fight in political struggle, it needs to do more than just provide Rawlsian 'reasonable faith'. They might want justice that is 'empirically likely', at least in the long run.

When it comes to feasibility, Rawls argues that "the fundamental principles of justice quite properly depend upon the natural facts about men in society" (Rawls, 1999a: 159). At the same time, Rawls claims "there are questions about how the limits of the practically possible are discerned and what the conditions of our social world in fact are. The problem here is that the limits of the possible are not given by the actual ... we have to rely on conjecture and speculation, arguing as best we can that the social world we envision is feasible and might actually exist, if not now then at some future time under happier circumstances" (Rawls, 1996: 12).

'Natural facts about men' and 'general facts of economics and psychology' are not necessarily facts in the same way that certain uncontroversial laws of physics or mathematics are. Even today, with much more sophisticated forms of scientific inquiry, we continue to dispute what can be considered a fact of human nature or economics. Secondly, even if we could find a set of uncontroversial facts, as Rawls suggests, these might not necessarily be constraints on our conceptions of justice. As stated before, what is possible is not limited by the actual (Rawls, 1996: 12). This suggests that we need to rely on 'conjecture and speculation' on a whole host of issues relevant to justice, including human motivation, economic efficiency, and what people might find politically

⁸⁶ Rawls provides a description of how this stability might emerge, starting with a constitutional consensus that eventually develops into an overlapping consensus. However, the discussion is very brief, short on details, and unlikely to satisfy anyone who believes that this is insufficiently realistic. (Rawls, 1996: 158-168)

intuitive. There is nothing wrong with this. However, it is easy to imagine how the people who find conviction from feasible theories will be divided amongst the theories that have different feasibility constraints. For example, someone might find conviction from Rawlsian justice and its reliance on a ‘sense of justice’ of reasonable citizens, while others might find that this conception of citizens is too idealistic for most (even reasonable) human beings and therefore find conviction from a theory that has a more ‘realistic’ account of human motivation. The same can be said for the ‘facts of economics and psychology’ that different theories of justice might incorporate. As Mason argues “It is hard to see why we should think that there is a uniquely correct account of the feasibility condition that the principles of justice defended within ideal theory should meet ... ideal theories may reasonably interpret the feasibility condition differently” (Mason, 2016: 52). This inevitably splits the base of people who might obtain a depth of conviction from Rawlsian theories of justice, which means that its ability to play the conviction role in political struggle is weakened; it is not able to generate the level of solidarity required for effective political movements.

Reconciliation

Whilst I am more sympathetic to Rawls’s theory of justice compared to Cohen’s, I believe the limitations of weaponising Rawlsian constructivism I discussed are significant enough that there is little reason to weaponise Rawlsian theory in the way that the conviction model wants. However, if we want to generate meaningful political impact, this does not mean that we can rely on Cohen’s justice alone. I think the way Cohen described the model does not acknowledge the fact that there are other ways in which we can generate depth of conviction, particularly if someone is reasonable. To keep things relatively ecumenical, I define reasonable people as are those who believe in the basic tenets of the public political culture of a liberal democratic society and are willing to sincerely engage in moral and political deliberation. Can Cohen’s approach to justice develop the depth of conviction required for reasonable people who are engaged in deliberation with other reasonable people?

One of the main issues with Cohen’s intuitionist approach is that for most political philosophers, providing the best account of one virtue, namely justice, is not going to be enough. I agree with Cohen that we have very deep “normative convictions, which being, our deepest, we find it difficult to defend (except against attack)” (Cohen, 2008: 4). However, there seems to be something incomplete about saying “Justice is not the only value that calls for (appropriately balanced) implementation: other principles, sometimes competing with justice, must also be variously pursued and honoured”, whilst not providing a clear account of how exactly we balance justice with other values when we ask the question ‘what should the state do?’ and ‘what rules should we live by?’ (Cohen, 2008: 271-272). Reasonable individuals will not find that to be enough for continuing in effective political

struggle. Any conviction will be undermined when we inevitably face difficult choices that require trade-offs of different values. Reasonable people will want to know that they can make the appropriate trade-offs and reach those conclusions in a principled manner.

Rawls's aspiration is "to devise a more systematic theory of distributive justice that economises on appeal to intuitive judgments about how to balance competing values" (Williams, 2013: 83). For most reasonable people, "Cohen's position is unsatisfying because it does so little to meet political decision-makers' need for some principled verdictive guidance, and rests content to express regret at the presence of an alleged distributive imperfection" (Williams, 2013: 83). To provide the depth of conviction for reasonable people to deliberate with fellow reasonable individuals before exposing themselves to the rough and tumble of political struggle, our theories of justice need to do more than providing a beautiful picture of the world. We need to do more than just saying that a state of affairs is "in one way better" even though all things considered, we should probably not level down (Cohen, 2011: 231).

Is there a way to reconcile these two models of theory and practice in satisfying manner? I believe there is. I believe that Rawlsian ideal theory, which provides the systematic theory to balance competing values, should (ideally) be accompanied by a radical Cohenite theory of ideals for effective political deliberation.⁸⁷ This combination allows for a depth of conviction for reasonable people who want a systematic account of our considered judgments of the just and unjust when engaging in moral deliberation. This can provide politicians and political agents who need to make difficult decisions with the depth of conviction that they are making principled trade-offs. The radical Cohenite theory of ideals can then be an effective tool for politicians and political activists to engage in arduous political struggle.⁸⁸

As mentioned in the introduction, "the purpose [of a theory of ideals] is to identify, elucidate, and clarify the nature of an ideal or ideals" (Hamlin and Stemplowska, 2012: 53). Many works in the methodological debates

⁸⁷ As stated before, my argument applies to theories other than Rawls's.

⁸⁸ The same can be said about non-Rawlsian theories of justice that have similar limitations.

in political philosophy have cited the distinction set out by Hamlin and Stemplowska. However, few have developed the idea of a theory of ideals much further. This lack of development is understandable. Not only did Hamlin and Stemplowska provide a thorough analysis, I think most people would presume that a (partial) theory of ideals is necessary to engage in any type of normative theorising. All philosophers want to provide a plausible definition of a value before engaging in normative analysis, which attempts to balance that value with others. The conviction model of theory and practice shows us that a theory of ideals is not just a necessary step towards an ideal theory. A theory of ideals can actually be a weapon for political agents engaging in ‘real politics’.

If we think about the Cohenite conviction model of theory and practice, asking ‘what is the distributive role of justice?’ without complicating alien factors and without the aim of providing all things considered judgment about our fundamental terms of association can have radical implications for political agents engaging in political struggle.⁸⁹ The simplicity of that answer can generate effective justificatory rhetoric to push for (or resist) change. An answer that represents our ultimate convictions can inspire solidarity amongst politicians and activists to continue their struggle even when times are difficult. It can also inspire political agents to demand the impossible with the aim of being realistic. Demanding the impossible is about inspiration rather than providing a feasible roadmap. Finally, the ideological self-confidence can ensure egalitarian political agents fight for incremental progress towards equality without doing the work of the Right for them. Politicians and political activists can generate conviction and solidarity amongst a wider range of the population, without having to rely on citizens to read (and comprehend) every aspect of a complex normative theory.

⁸⁹ Rawlsians generally believe that Cohen did not have a compelling enough answer to the question ‘what’s in a name?’ (Pogge, 2008). Is there a significant cost to Rawlsians accepting Cohen’s arguments and relabelling their work as theories of ‘rules of regulation’? We want disputes in political philosophy to be normative rather than linguistic, and any victory that is merely linguistic is unsatisfying. The conviction model of theory and practice provides one compelling Cohenite answer to the question ‘what’s in a name?’ Answering the question ‘what is justice’ in the way that Cohen wants can have radical political implications.

Conclusion

Throughout this paper, I have broadly focused on a tension that exists between different types of theory. A theory can provide us with systematic guidance, but lose out on its ability to inspire political agents to fight in the arduous struggle of politics, or it can inspire political agents, but provide us with little to no guidance on how to properly balance values and how to make principled trade-offs. However, this tension need not be an issue if we accept that both Rawlsian and Cohenite approaches to justice can play very different roles. When it comes to converting theory into practice, we should try our best to ensure that the Rawlsian approach is complemented with the Cohenite conviction model to help political agents generate political impact. The relationship between these two conceptions of the relationship between theory and practice should also show that non-ideal circumstances does not necessarily mean we need to ‘water down’ our ideals into a non-ideal theory. Real world circumstances can give us reasons to demand the impossible.

References

- Anderson E (2010) *The Imperative of Integration*. Princeton: Princeton University Press
- Cohen G. A (2008) *Rescuing Justice and Inequality*. Cambridge: Harvard University Press
- Cohen G. A (2009) *Why Not Socialism?* Princeton: Princeton University Press
- Cohen G. A (2011) *On the Currency of Egalitarian Justice*. Princeton: Princeton University Press
- Hamlin A and Stemplowska Z (2012) Theory, Ideal Theory and the Theory of Ideals. *Political Studies Review*, 10(1): 48-62
- Mason A (2016) Justice, Feasibility, and Ideal Theory: A Pluralist Approach. *Social Philosophy and Policy*, 33: 32-54
- Miller D (2013) *Justice for Earthlings: Essays in Political Philosophy*. United Kingdom: Cambridge University Press
- Murphy L and Nagel T (2002) *The Myth of Ownership*. New York: Oxford University Press
- Nozick R (1974) *Anarchy, State and Utopia*. USA: Basic Books
- Pogge T (2008) Cohen to the Rescue!, *Ratio*, 6: 88-10
- Rawls J (1996) *Political Liberalism*. 2nd ed. New York: Columbia University Press
- Rawls J (1999a) rev ed. *A Theory of Justice*. Massachusetts: Harvard University Press
- Rawls J (1999b) *Collected Papers*. Cambridge: Harvard University Press

Rossi E and Sleat M (2014) Realism in Normative Political Theory. *Philosophy Compass*. 9(10): 689-701

Sen A (2009) *The Idea of Justice*. USA: Harvard University Press

Stemplowska Z and Swift A (2017) Ideal and Nonideal Theory. In: Estlund D (ed). *Oxford Handbook of Political Philosophy*. New York: Oxford University Press

Swift A and White S (2008) Political Theory, Social Science, and Real Politics. In: Leopold D and Stears M (ed). *Political Theory: Methods and Approaches*. New York: Oxford University Press

Williams A (1998) Incentives, Inequality, and Publicity. *Philosophy and Public Affairs*. 27(3), 225-247

Williams A (2008) Justice, Incentives and Constructivism. *Ratio*, 7: 110-126

Williams A (2013) How gifts and gambles preserve justice. *Economics and Philosophy*, 29: 65-85

6. Ideal and Systematic Theory

WORKING PAPER by Aaron Landau (Columbia University)

The practical motivation for constructing an “ideal theory” of a perfectly just society is that it is essential for justifying some of the principles for “nonideal theory,” which tells us how to act when there is substantial injustice. As John Rawls says, “[t]he reason for beginning with ideal theory is that it provides ... the *only basis* for a *systematic grasp* of [the] more pressing problems” of nonideal theory.⁹⁰ Among the various arguments for this “priority” or “dependence thesis,” two are especially prominent. First, ideal theory is necessary for the justification of priority rules (e.g. Rawls’s famous lexical ordering) that show us which injustices to remove before others. Second, we must do ideal theory because we need to know how we should transition to a fully just society.⁹¹ In this paper, I argue that both these claims are mistaken and that we can establish first principles of justice and construct a systematic nonideal theory independently of knowledge of the ideal.

1. Ideal Theory, Ideal Principles and Nonideal Theory

⁹⁰ John Rawls, *A Theory of Justice*, Rev. ed (Cambridge, Mass: Belknap Press of Harvard University Press, 1999), 8 emphasis added.

⁹¹ Rawls, 216, 343; John Rawls, *The Law of Peoples with “The Idea of Public Reason Revisited”* (Cambridge, Mass: Harvard University Press, 1999), 89; A. John Simmons, “Ideal and Nonideal Theory,” *Philosophy & Public Affairs* 38, no. 1 (2010): 5–36; Zofia Stemplowska and Adam Swift, “Rawls on Ideal and Nonideal Theory,” in *A Companion to Rawls*, ed. Jonndle and David A. Reidy (John Wiley & Sons, Inc., 2013), 112–27; Ben Laurence, “Constructivism, Strict Compliance, and Realistic Utopianism,” *Philosophy and Phenomenological Research*, 2017, Early View.

The concept of “ideal theory” is understood in several different ways in the literature, but I will focus only on two core definitions.⁹² The first is *strict compliance theory*, which justifies principles of justice under the assumption that “(nearly) everyone strictly complies ... with [those] principles.”⁹³ In contrast, *partial compliance theory* establishes the principles that are suitable for conditions in which there is substantial noncompliance. The second definition is *realistically utopian theory*, which justifies principles of justice on the assumption that the society is a “realistic utopia” or one that is *as just* as is “achievable” in our social world, such that it “is feasible and might actually exist, if not now then at some future time under happier circumstances.”⁹⁴

When I talk about a perfectly just, a just, or an ideal society, I mean one that is either strictly compliant or realistically utopian. Still, there is liable to be some misunderstanding about what exactly counts as an ideal theory, which I hope to prevent by offering the following definition and explication of it.

Ideal Theory: A theory that contains an argument with the explicit aim of justifying principles of justice for a society in which (nearly) everyone strictly complies with those principles or for a realistic utopia.

A theory is an ideal theory, then, if and only if it has the explicit aim of justifying principles for these very specific social conditions. These are accounts that are intentionally constructed to answer the question: What principles should be followed in a strictly compliant or realistically utopian society? This means that an ideal theory will essentially include an argument that reasons from empirical premises stating that the relevant society is specifically a strictly compliant or a realistically utopian one. Similarly, I propose that we should say that a

⁹² See, for example, Laura Valentini, “Ideal vs. Non-Ideal Theory: A Conceptual Map,” *Philosophy Compass* 7, no. 9 (2012): 654–664.

⁹³ John Rawls, *Justice as Fairness: A Restatement*, ed. Erin Kelly (Cambridge, Mass: Harvard University Press, 2001), 13.

⁹⁴ Rawls, *The Law of Peoples*, 6, 12.

principle is an “ideal principle” if and only if it is the conclusion of an argument with the explicit goal of justifying principles for a perfectly just society.

These distinctions are absolutely crucial for our topic because we need to distinguish ideal theory and ideal principles from other kinds of theory and principles that are independently available to nonideal theory. Consider, for example, utilitarianism. The principle of utility states that the right action is the one that would maximize preference-satisfaction. Among the key considerations that typically support utilitarianism are the all-encompassing value of preference-satisfaction and a maximizing principle of value. Moreover, the principle of utility is a universally applicable moral standard. Incidentally, this means that the principle of utility determines what should be done in a just society. However, utilitarianism is not an ideal theory and the principle of utility is not an ideal principle because their supporting arguments are not designed to ascertain the principles of a strictly compliant or realistically utopian situation in specific.⁹⁵ I will examine some non-consequentialist theories later, but the general point is that there are theories and principles that are not ideal ones. Therefore, if ideal theory and ideal principles are to be essential for nonideal theory, they have to make a contribution that these other kinds of theory and principles cannot.

In contrast to ideal theory, *nonideal theory* justifies principles of justice for circumstances that are less just than a strictly compliant or realistically utopian society. I also stipulate that the salient kinds of nonideal theory are only those about *contemporary* unjust circumstances, that is, present situations or those in the not too distant future. This is to ensure that the kinds of nonideal theory we discuss are practically significant for us.

2. The Dependence Thesis

⁹⁵ This is statement about utilitarianism as it is canonically justified. Of course, we may explicitly ask (as Rawls does) whether the principle of utility should be adopted in a strictly compliant society, in which case, we would be considering the principle of utility as an ideal principle. Moreover, we can do ideal theory with utilitarianism by explicitly asking what the principle of utility entails for a just society. The point is that utilitarianism and the principle of utility are not essentially or usually an ideal theory and an ideal principle. Rawls, *A Theory of Justice*, secs. 27–28.

Another basic concern is how to characterize “the dependence thesis” itself.⁹⁶ I propose that the following formulation is suitable for each of the more specific theses we will examine.

Dependence Thesis: The completion of ideal theory is a necessary condition for justifying some of the principles that would better facilitate making sound judgments about how to respond to the problems of injustice and nonideal theory.⁹⁷

One way of putting the same idea is that ideal theory is required for the construction of a more “systematic” and “complete” nonideal theory of justice. The resulting theory is more *systematic* in the sense that it allows us to identify and to explain which judgments are right in terms of more fundamental principles of justice. In addition, the theory is more *complete* in that its principles would entail or help support the right judgments for a broader set of problems of injustice.⁹⁸

3. First Principles of Justice

Tommie Shelby argues that ideal theory is essential for providing “a philosophically adequate and complete theory of how to respond to social injustice” because “ideal principles” are the most fundamental or first

⁹⁶ I use the term “dependence thesis” instead of the more common “priority thesis” in order to avoid confusion with what Rawls calls “priority rules.”

⁹⁷ Although I will not engage in an exegetical argument, I believe this formulation expresses the core idea of Rawls’s view that “[t]he reason for beginning with ideal theory is that it provides ... the *only basis* for a *systematic grasp* of [the] more pressing problems” of nonideal theory. Ultimately, whether my definition of the dependence thesis is appropriate is to be judged by whether it captures the shared claim of the more specific theses in the literature. Rawls, *A Theory of Justice*, 8.

⁹⁸ In part, my characterization of the dependence thesis stems from Rawls’s discussion of “the full system of principles” or “full conception of right” and its “completeness.” Rawls, 299–301.

principles of justice that explain “what makes a social scheme unjust” and “what direction social change should take.”⁹⁹ Let us call this the “first principles thesis.” However, we have already seen that these claims are not generally true. The basic principle of utilitarianism is not an ideal principle, and it can be used in nonideal theory to ascertain the right courses of action by determining which would maximize (expected) utility. For utilitarianism, then, fundamental principles for identifying injustice and right action in nonideal theory do not depend on ideal theory.

It might be objected that utilitarianism can avoid ideal theory because it is a version of consequentialism. Perhaps, ideal theory is indispensable for some deontological theories, especially those in which equality or fairness is a basic value. However, a structurally identical critique also applies to contractualism. At the root of T.M. Scanlon’s contractualism is a certain view of what is morally valuable about persons, that is, that they are rational agents with the capacity to assess and act on reasons or justifications. Respecting the value of persons requires complying with the following first principle of interpersonal morality: we are to treat others in accordance with principles that are justifiable to each person or that “no one could reasonably reject.”¹⁰⁰ We assess the justifiability of a principle by comparing the burdens different feasible principles would impose on various *single* individuals in the salient circumstances. An individual can reasonably reject a principle just in case she can propose an alternative principle which is not comparably burdensome for any other individual in the circumstances. For any moral agent, then, the principles that are reasonably rejectable by none are those that generate (roughly) the weakest individual burden or complaint in the circumstances.¹⁰¹ Importantly, ideal theory is not a necessary part of the justification for the fundamental contractualist principle. Moreover, both contractualism’s basic principle and justificatory method can be directly applied to nonideal situations without

⁹⁹ Tommie Shelby, “Racial Realities and Corrective Justice: A Reply to Charles Mills,” *Critical Philosophy of Race* 1, no. 2 (2013): 156; Tommie Shelby, *Dark Ghettos: Injustice, Dissent, and Reform* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 2016), 11–13.

¹⁰⁰ T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass: Belknap Press of Harvard University Press, 2000), 153.

¹⁰¹ Comparable burdens, greater burden test vs slightly worse burdens etc.

taking a detour to ideal theory. In nonideal theory, contractualism justifies principles for any given agent based on the relative costs various principles would have on individuals in *unjust circumstances*; no input about the ideal is necessary. Hence, as with utilitarianism, ideal theory is not essential for deriving or applying the first principles that identify injustice or just social reforms in nonideal theory.

Proponents of the dependence thesis often assume that first principles of justice are equivalent to ideal principles.¹⁰² However, as we have seen, this presumption is not warranted. Indeed, although time does not permit it, I believe examining the justificatory structure of many other theories (including Rawls's) would also reveal that having first principles for nonideal theory does not require ideal theory. If, therefore, there is some significant class of theories that does need ideal theory to justify its first principles, it is up to defenders of the first principles thesis to produce the necessary argument, which they have not.

4. Priority Rules

One of the main ways in which ideal theory is taken to be essential for nonideal theory is by providing “priority rules” that rank or weigh the relative importance of different principles of justice. Call this “the priority rules thesis.”¹⁰³ In order to see that the priority rules claim is mistaken, we only need to modestly extend the arguments from the previous section. Utilitarianism would determine the relative importance of different principles just as it does actions, that is, in terms of how much they would increase overall preference-satisfaction. In some circumstances, satisfying some principles (e.g. implement the social conditions for personal autonomy) might generate more utility than meeting others. In that case, utilitarianism would imply priority rules

¹⁰² Rawls, *A Theory of Justice*; Laura Valentini, “On the Apparent Paradox of Ideal Theory,” *Journal of Political Philosophy* 17, no. 3 (2009): 332–355; Simmons, “Ideal and Nonideal Theory”; Shelby, *Dark Ghettos*; Hillel Steiner, “Debate: Levels of Non-Ideality,” *Journal of Political Philosophy* 25, no. 3 (2017): 376–84.

¹⁰³ E.g. Rawls, *A Theory of Justice*, 216, 343; Simmons, “Ideal and Nonideal Theory”; Stemplowska and Swift, “Rawls on Ideal and Nonideal Theory.”

that weight the former principles more than the latter, in those circumstances. Since there is no reason why this could not be straightforwardly done for nonideal situations, knowledge of the ideal is extraneous.

Similarly, contractualism justifies principles in terms of their comparative burdens for single individuals. We can simply extend this procedure to the justification of higher-order principles like priority rules and compare the burdens that are attached to various weightings or rankings of certain first-order principles. The right set of priority rules would weigh or rank a principle more or higher just in case following it over the others would burden individuals less than complying with the others over it. We would be able to derive such rules if we had an account of individual's interests and the priorities among them. Since contractualism holds that the power of rational autonomy is what is most valuable about human beings, we would expect that it would count the promotion of that capacity as persons' highest or lexically prior interest. It follows that priority rules that sacrifice that interest for other ones would entail greater burdens for individuals than those that promote it over others. Therefore, contractualism would produce priority rules that give precedent to principles that require the effective establishment of the means necessary to protect and cultivate rational autonomy. If these means include the basic liberties (e.g. freedom of conscience, freedom of association, etc.), then we would have a contractualist argument for what Rawls calls "the priority of liberty."¹⁰⁴ However, since the basic liberties are of first importance for individuals' interests in many nonideal circumstances and the contractualist method can be applied directly to these situations, the priority of liberty does not require ideal theory for its justification.

As with first principles, I believe that ideal theory is not necessary for justifying priority rules in other moral frameworks as well (including Rawls's). Given the above counterexamples and lacking any argument to the contrary, we do not have good reason to think the priority rules thesis is credible.

5. Realistic Utopia as a Practical End

¹⁰⁴ Rawls, *A Theory of Justice*, sec. 82.

The most prominent dependence claim in recent times is “the realistic utopia thesis” that ideal theory is necessary for nonideal theory because we need to know how to best transition to our ultimate goal of a realistic utopia.¹⁰⁵ However, defenders of the thesis often assume, rather than demonstrate, that obtaining perfect justice is an aim that is practically salient for us.¹⁰⁶ But, this is not obvious. For example, a just society may seem like an extremely improbable outcome, which we can safely ignore. It is important, therefore, to closely examine the reasoning that supports the realistic utopia thesis.

Here is an argument that can be reconstructed from Rawls. He says that “as far as circumstances permit, we have a natural duty [of justice] to remove any injustices” and “to assist in the establishment of just arrangements when they do not exist, at least when this can be done with little cost to ourselves.”¹⁰⁷ We may not be able to advance justice because it is beyond our abilities or the available means may be morally impermissible in the circumstances, and we may be permitted to tolerate injustice if remedying it is too burdensome. The probability of success is also a factor that affects whether we can or it is permissible to attempt to bring about more justice. However, if we are able, it is morally permissible, sufficiently likely, and not too costly, the natural duty implies that we ought to eliminate as much injustice as we can within these limits. Furthermore, it also means that, if the circumstances are favorable, we are to continue to rectify injustice until it is completely extinguished and a just society is realized.

¹⁰⁵ E.g. Simmons, “Ideal and Nonideal Theory”; Stemplowska and Swift, “Rawls on Ideal and Nonideal Theory”; Laurence, “Constructivism, Strict Compliance, and Realistic Utopianism.”

¹⁰⁶ For example, Simmons asserts that our final goal is to achieve a realistic utopia as a rejoinder to Sen’s criticisms of Rawls, but we have no reason to believe that Sen would see perfect justice as a practically relevant outcome. Simmons, “Ideal and Nonideal Theory,” 34ff.

¹⁰⁷ The natural duty of justice is a duty for individuals. However, we may also apply it to institutional and group agents too. It would not affect the argument. Rawls, *A Theory of Justice*, 216, 294.

I propose that we understand the major premise of the argument as that we are to take the actions with the most expected justice, if they are within our powers, are morally permissible, and are not too costly. To explain the notion of “the greatest expected justice,” consider the following case from Frank Jackson:

Drug: Jill is a physician who has to decide on the correct treatment for her patient, John, who has a minor but not trivial skin complaint. She has three drugs to choose from: ... Careful consideration of the literature has led her to the following opinions. Drug A is very likely to relieve the condition but will not completely cure it. One of drugs B and C will completely cure the skin condition; the other though will kill the patient, and there is no way that she can tell which of the two is the perfect cure and which the killer drug. What should Jill do?¹⁰⁸

The special feature of cases of this kind is that it is only permissible for the moral agent to choose an action that her evidence indicates will not lead to the best outcome of those in fact available. The best outcome is a complete cure, which would be the result of administering either drug B or C. Nevertheless, in light of the risk that treating John with B or C also has an equal chance of killing him, it would be wrong to give him either. Therefore, even though there is a lower probability that drug A will restore John to perfect health than either B or C, Jill has a moral duty to choose A. What Drug strikingly shows is that our moral duties, at least in terms of consequences, are not determined by what actions would *in fact* lead to the best outcome but by what the *evidence* indicates is the action that provides the *best chance* of achieving what is of value and of avoiding what is of disvalue. One familiar way of specifying this idea is that the morally right action is the one that *maximizes expected value*, where expected value is a function of the probabilities of the action’s possible outcomes and the actual values of those outcomes. Acknowledging that various qualifications may be appropriate, let us say that, all things being

¹⁰⁸ Frank Jackson, “Decision-Theoretic Consequentialism and the Nearest and Dearest Objection,” *Ethics* 101, no. 3 (1991): 462–63.

equal, our moral duty is to take the action that has the greatest expected justice, i.e. the best chance of attaining justice and avoiding injustice.

It is reasonable to suppose that the requirement of greatest expected justice is implicit in Rawls's argument because the natural duty of justice demands the removal of "any injustice," i.e. gaining as much justice as possible if the circumstances permit, and because he recognizes that the likelihood that actions will effectively improve justice is a relevant consideration. The proposal, then, is that the natural duty of justice says that we are to take the actions that maximize expected justice in our circumstances, all things being equal.¹⁰⁹

Since it is based on the natural duty of justice, the argument for the realistic utopia thesis must be that, in some way, taking into account the prospect of achieving a realistic utopia will help us ascertain which actions have the greatest expected justice in our circumstances. It follows that if, in our circumstances, the outcome of a just society cannot make a difference to identifying the actions that maximize expected justice, then the thesis is unjustified.

We have seen that the expected justice of an action is a function of both the probabilities of its outcomes and their actual values as indicated by the evidence available to us. Since the value of an outcome is independent of our actions, a prospect can affect the relative expected justice of different actions only if we can ascertain, from our evidence, how the probability of its occurrence would change if those various actions were to be taken. In other words, an outcome is significant to whether we should choose one action over another only if we have the ability to predict how these actions will make its realization more or less probable. This means that a social prospect cannot influence which action maximizes expected justice in nonideal theory if we cannot determine what its likelihood would be conditional on our available actions. The problem for the prospect of a realistic utopia is that we simply are not able to predict how our actions will influence its probability based on the

¹⁰⁹ For the sake of simplicity, I will proceed on the assumption that the actions in question are morally permissible and not exceedingly costly. This is acceptable because the relevant argument is that realistically utopian theory is essential for identifying more just *consequences*. It is possible to criticize the realistic utopia thesis using considerations of permissibility and over-demandingness, but that will not be necessary.

evidence available to us. This is because of the facts that the realization of complete justice would involve radical social changes from the current state of affairs that would take a very long time to accomplish and that our social systems contain highly complex causal processes that interact with one another and try to influence the world in conflicting ways. Therefore, in order for contemporary agents to reliably predict what difference, if any, their actions would make on the likelihood of perfect justice, they would need extensive empirical knowledge of the economic, social, political, ideological, national, international, technological, psychological, etc., facts and trends of the world. In addition, they must be able to predict how these various kinds of facts do and will interact and how the actions available to us would affect this causal nexus as it relates to the probability of a fully just society being achieved in the distant future. It is safe to say that no one has the knowledge or the cognitive powers to make the necessary predictions. It follows that knowledge of a realistic utopia and ideal theory cannot assist us in maximizing expected justice in nonideal theory. The argument for realistic utopia thesis, therefore, fails.

Conclusion

There are other senses in which ideal theory is said to be necessary for nonideal theory, which I examine elsewhere.¹¹⁰ Nevertheless, it is important that the most influential arguments for the dependence thesis are not successful. A systematic nonideal theory does not need ideal theory in order to justify first principles of justice or priority rules or to maximize expected justice in our circumstances.

¹¹⁰ Other versions of the dependence thesis include the claims that ideal theory is essential for evaluating a society's overall justice and injustice, for discerning whether the state or public has failed in its duties, and for identifying our fair shares or level of sacrifice in nonideal theory. I address all of these in the full version of this paper. Rawls, *A Theory of Justice*, 216; Shelby, *Dark Ghettos*; Liam B. Murphy, *Moral Demands in Nonideal Theory* (Oxford: Oxford University Press, 2003).

