

# *Norms, evaluations and ideal and non-ideal theory*

Article

Accepted Version

Jubb, R. (2016) Norms, evaluations and ideal and non-ideal theory. *Social Philosophy and Policy*, 33 (1-2). pp. 393-412. ISSN 1471-6437 doi:  
<https://doi.org/10.1017/S0265052516000212> Available at  
<https://centaur.reading.ac.uk/67057/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1017/S0265052516000212>

Publisher: Cambridge University Press

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

## Norms, Evaluations and Ideal and Non-Ideal Theory

### Introduction

The debate in political theory and philosophy about ideal theory has now been going in its current form for around a decade. Amartya Sen's 2006 *Journal of Philosophy* piece, 'What Do We Want from a Theory of Justice?', seems to be its most widely shared starting point (Sen 2006). Sen was certainly not the first person to worry about contemporary political philosophy's tendency to ignore various recalcitrant features of the real world. One could sensibly think of much of the communitarian and feminist critique of Rawlsian liberalism in the two decades following the publication of *A Theory of Justice* in this way, for example. However, Sen's critique of the "transcendental approach" recast and reinvigorated debates around idealization and abstraction. His insistence that we only need comparative evaluations to identify both injustices like "the persistence of endemic hunger or exclusion from medical access" and what would remedy them has often been taken since as the most important recent attack on ideal theory (Sen 2006: 224).

For example, in Adam Swift's contribution to the special issue of *Social Theory and Practice* on ideal theory he and Ingrid Robeyns edited, Sen is the first named critic of ideal theory and a central focus of much of the rest of the article (Swift 2008: 365ff). After some initial clarificatory work, Swift is primarily concerned to vindicate ideal theory against Sen's observation that "a full specification of "spotless justice" is neither necessary nor sufficient to... guide action in the circumstances that actually confront us" (Swift 2008: 365). And this is not only a feature of discussions immediately following Sen's piece. For instance, David Wiens describes his recent attack on the capacity of ideal theory to provide guidance as an attempt to "go beyond Sen's arguments to show that political ideals are misleading or uninformative... in nonideal circumstances" (Wiens 2015: 435). For Wiens, it is Sen and not any of the many other, otherwise similar, critics of ideal theory whose arguments do not quite find their targets and so need to be put right.

Of course, the debate Sen started has fractured and fed into or overlapped with various other discussions which are not always properly distinguished. Laura Valentini's typology of discussions of ideal theory separates three different disputes, over full and partial compliance, utopianism and realism, and end-state and transitional theory (Valentini 2012). Yet that typology is misleading. For one thing, it mistakenly assimilates realist dissatisfactions with the utopian moralism of contemporary political philosophy to the kind of complaint Sen made (Jubb 2015b: 919; see also Rossi and Sleat 2014: 690). Realists' most widely shared initial reference point seems to be Bernard Williams' 'Realism and Moralism in Political Theory' (see for instance Galston 2010, where Williams' piece is "the best entry point into [its] topic": 387). That is at

least as hostile to Sen's demand for comparisons between different states of affairs as it is to Rawls' transcendental theorising. Williams begins with a pair of contrasts, one of which overlays the other. First, he contrasts "enactment" and "structural" models of the relation between morality and political practice (Williams 2005: 1). He then observes that they "both represent the priority of the moral over the political" and so are forms of moralism. While a structural model sees politics as requiring moral "constraints", an enactment model treats politics as an "instrument of the moral" (Williams 2005: 2). Both can therefore be contrasted with Williams' preferred alternative, realism, which instead "gives a greater autonomy to distinctively political thought" (Williams 2005: 3). Sen's view is clearly an enactment view. Its demand that theories of justice identify morally urgent and achievable social improvements makes politics into an 'instrument of the moral' in exactly the way that Williams treats as definitive of an enactment view. So realism is not helpfully thought of as about ideal or nonideal theory, at least insofar as that discussion starts with Sen.

The confusion of the debates taking Sen as their starting point makes it very important to be clear about what the topic is when discussing ideal theory. My topic here is the relation between ideal theory and what Williams calls structural and enactment views. I engage that topic by discussing the way feasibility has been used in defences of ideal theory. Defending ideal theory by simply adding feasibility to it is, I try to argue here, a mistake. It is a mistake because it tends to replicate the justificatory structure which made ideal theory pointless or unintelligible to its critics. I use David Wiens' excellent work on the structure of a political theory to show this. This reveals an important feature of ideal theory, at least as I understand it here. Ideal theory makes sense only for a structural view, which focuses on constraints and so on the agents subject to them. Enactment views, which focus on improvements and so the states of affairs which realise them, will instead find ideal theory useless. My main aim in this paper is to suggest this connection between structural and enactment views and the interest of ideal theory. Observing this connection will help turn aside a series of effectively question-begging critiques of ideal theory, at least if structural views pass a bar of basic plausibility. That will clarify the confused terms on which the debate has been taking place and so allow it to progress to more important questions like how we ought to understand content of our political ideals. For example, how should we understand contemporary commitments to equality and their relation to the demands for legitimacy that stable representative democracies must met (Jubb 2015a)?

It is also important to be clear about how terms are being used in a discussion of ideal theory. I will use ideal theory to mean a more or less complete specification of the institutional and individual rights, duties, privileges, powers and immunities which together characterize some suitably perfected society or other set of relations. This is what Sen seems to mean by a

transcendental theory, which he describes as attempting to identify “perfectly just societal arrangements” (2006: 216). That in turn is close enough to Rawls’ initial use of the term, which similarly focuses on the idea of perfection (Rawls 1971: 8-9). My usage therefore should cover much of what is meant both by those who criticize ideal theory and their primary target. This is not how the term, ideal theory, is always used. For example, Zofia Stemplowska’s defence of ideal theory treats any theorizing which does not produce achievable and desirable recommendations as ideal theory (Stemplowska 2008: 324ff). This obviously captures much more than ideal theory in my sense, including theorizing about situations that are much worse than ours and which consequently does not seem appropriately described as ideal at all (see Jubb 2012 for more detailed arguments against this understanding of ideal theory). In comparison to Stemplowska, I am interested in a much narrower range of abstractions and idealizations. However, we have already seen that there are various different discussions going on under the heading of ideal theory. Defining your terms will inevitably mean excluding some of them. Not doing so means at best talking at cross purposes.

I will also rely on a contrast between norms and evaluations. This distinction tracks a distinction similar to that which Williams’ contrast between structural and enactment views aims to map. Norms demand actions and evaluations assess or rank objects, whether concrete, like paintings, or abstract, like states of affairs. The emphasis structural views place on constraints makes them primarily normative. Enactment views contrastingly typically emphasize evaluations since they are needed to identify the improvements at which they aim. Of course, particular norms may demand that certain states of affairs are brought about, just as the assessment or ranking of states of affairs may depend on which actions have been performed. That norms and evaluations may relate to each other however does not mean they are reducible to each other. Norms can only apply to agents and no evaluation is rationally binding, either in thought or (other forms of) action, in the absence of a norm. For example, a painting may constitute a violation of a norm, perhaps because its creator promised not to paint that subject, but it cannot violate a norm itself because it cannot act. And the evaluation, that painting is bad, can only properly compel anyone in conjunction with a norm. Absent a norm of some sort, it has no purchase on agents. It cannot even require belief in itself. Only a norm about when to believe given propositions can require belief in anything.

The rest of this paper proceeds as follows. In the first section, I recapitulate the beginnings of the current debate over ideal theory. I try to show that Sen’s article displays a commitment to the priority of evaluations over norms. Claims in political theory are supposed to direct us how to make the world better, and are pointless if they do not. I also claim that two other early attacks on ideal theory also focus on producing higher-ranked states of affairs. In the second

section, I go on to try to show that one common response to that critique of ideal theory often assumes much the same kind of priority of evaluations over norms. This response accepts that it is problematic that ideal theory does not straightforwardly provide instructions for would-be reformers. It tries to put that right by adding feasibility to ideal theory and so producing a ranking of the options actually available, often urging that we pick whichever maximizes expected value. That, I claim, invites a range of anti-consequentialist criticisms. As David Wiens shows, it also fails on its own terms. The third and final section argues that a commitment to the priority of evaluations over norms tends to systematically miss the point of ideal theorizing. In the sense that I use it, ideal theory describes situations in which a particular set of relations or practices are appropriately governed. Governance is a normative, not an evaluative, notion. It is achieved if the relevant agents comply with various rules. Why would we expect basically evaluative perspectives to grasp the interest of such an achievement? I also argue that focusing on evaluating options may not be an improvement on using ideals to guide our political choices. Ranking alternatives may be just as difficult as trying to understand what acting rightly in a wrongful world would involve.

### **Recipes for Cook-Shops of the Future: Ideal Theory's Failure to Prescribe**

Sen's 'What Do We Want from a Theory of Justice?' advances a number of critiques of ideal theory. However, his claim that ideal theory is 'neither necessary nor sufficient for answering questions on the advancement of justice that urgently demand our attention' is central (Sen 2006: 237). His emphasis on the inclusive and emancipatory possibilities opened up by acknowledging the incompleteness of many of our judgments of justice would be irrelevant without it. If theories of justice had to be complete to generate answers to urgent questions, then what we could do when they were incomplete would hardly matter.

Sen does not claim that theories of justice which do not answer those urgent questions are no longer normative and therefore false, as for example Charles Mills and Colin Farrelly do (Mills 2005: 171; Farrelly 2007: 845). That is to Sen's benefit, since it is uncomplicatedly incorrect to claim that a theory of justice which does not tell us what to do now does not guide action and is therefore false. A theory of justice for beings quite unlike us may well be correct as far as it goes. It can certainly provide instructions appropriate for beings like that, were they around to receive them. That a prescription does not guide our action here and now does not mean that it is not action guiding. To deny this is effectively to deny that a conditional can be true when its antecedent is false. Mills and Farrelly have no reason to claim that normative theories for agents

importantly different from as such fail unless they also think that the statement, if the switch is flicked the light will come on, is always false if the switch is not flicked.<sup>1</sup>

However Sen does share something important with Mills and Farrelly. All three are committed to the priority of evaluations over norms. All three assess whether a theory of justice provides useful guidance in terms of its ability to identify accessible improvements to our current state of affairs. For Mills and Farrelly, this is straightforward. On Mills' account, theories which do not guide our action now and so "make ourselves better people and the world a better place" break the "link with practical reason" (2005: 170, 171). Similarly, Farrelly claims that if trying to follow a theory of justice "would not result in any noticeable increase in the justness of one's society, then it fails as a normative theory" (Farrelly 2007: 845: italics suppressed). For both, guiding action appropriately means directing it to improve ourselves, our world or society. The normative content of a theory must be given by a set of specific local evaluations which identify such improvements. Sen's comparative approach likewise "concentrate[s] on ranking alternative societal arrangements". Such an approach can show that "the introduction of social policies that abolish slavery, eliminate widespread hunger, or remove rampant illiteracy... yield an advancement of justice" (2006: 216, 217). For Sen, "answering questions on the advancement of justice that urgently demand our attention" requires evaluation. In contrast, an approach which focuses on specifying perfectly fulfilled norms cannot respond to the "extraordinary" injustice of our world (2006: 237). The content of practically relevant norms is given by the evaluations used to produce rankings of states of affairs.

This priority could be merely semantic. The evaluations which Sen, Mills and Farrelly insist should provide the content of norms could themselves refer to norms. Perhaps when Mills insists that a theory or principle is not normative unless it makes the world better, that is compatible with the world being made better by having more promises (rightly) kept. That does not obviously seem to be the case with Farrelly though. He argues that theorising about rights should be replaced with cost-benefit analysis (2007: 848ff). That argument urges the replacement of a set of norms governing agents with ranking of states of affairs by their net individual interest satisfaction. His commitment to treating politics as concerned with objects rather than actions is then not merely superficial.

---

<sup>1</sup> There are three things this argument does not show that might have led Farrelly and Mills into the mistake of claiming that theories which do not provide presently helpful guidance are false. First, it does not show that theories which do not give us instructions about what to do now are useful in some other way. There are though plenty of pointless truths. Second, it does not show that any theories are conditional in the relevant sense. Farrelly's and Mills' claims are however suitably general. Third, it does not show that we have the epistemic tools to assess or justify theories dealing with circumstances significantly different from our own. But truths do not have to be accessible to us.

Sen's discussion seems to reflect similar priorities. We can see this in two ways. First, Sen judges the satisfactoriness of norms according to the advancement of justice they achieve. The urgency of removing injustice is what condemns transcendental theory. It is unsatisfactory because it does not advance justice. Norms do not just have their content given by evaluations. The same evaluations also decide whether they are appropriate. Sen may not bind the truth of norms to their production of improvements in the way Mills and Farrelly appear to, but he does seem to judge the rightfulness of articulating, defending and discussing them according to the 'advancements of justice' they achieve. Second, Sen's discussion of the demands of justice is often strikingly agentless. Who achieves the advancements of justice to which he refers? A social policy is not itself an agent. Equally, the requirements of justice Sen mentions are typically not actions or omissions agents must perform. Being fed or literate is a state rather than something an agent does. Even when Sen's justice does constrain agents, it constrains them not to treat others as patients by enslaving or torturing them (2006: 218). There is nothing like the demand of reciprocity between agents *qua* agents Rawls' difference principle embodies (see for example Rawls 1971: 102ff). Sen's judgments of the appropriateness of theories of justice as well as of their content demonstrate the priority he gives to an evaluative over a normative perspective. His commitment to what Williams calls an enactment rather than a structural model of political moralism and so to the priority of evaluations over norms seems deep.

When Williams distinguishes enactment and structural models of moralism, he uses Rawls' theory to exemplify and so illustrate what he means by a structural view (2005: 1). In *A Theory of Justice* Rawls certainly meant to offer an alternative to utilitarianism, which Williams uses as an exemplar of an enactment model. And his method was part of that alternative. If Sen and others are as committed to political philosophy being primarily evaluative as I have tried to show, then it should not be surprising that they are hostile to Rawls' method. They structure their theories around evaluations in a substantive way which Rawls' insistence on the priority of the right over the good, for example, is meant to reject (1971: 29ff). Rawls' ideal theory is part of that method. It focuses on perfection and so on full compliance with the norms he prescribes. It hopes to show what agents must do, politically, to fully respect each other as free and equal. The only evaluation it is meant to produce, that a society at the specified level of development fulfilling the two principles is as just as it can be, is derivative of the fulfilment of norms. Sen's hostility is then predictable. Norms should be derivative of evaluations, not the other way round. Rawls' theory would have failed on its own terms if it was successful on Sen's.

### **Recipes for Cook-Shops Now: Feasible Prescriptions**

Defences of ideal theory against this kind of critique have come in two varieties. The first of these resists Sen, Mills and Farrelly by urging that ideal theory can be useful in non-ideal circumstances. A. John Simmons notes the attacks of all three, and taking Sen as their representative, argues that they have underestimated the importance of ideal theory. Sen has missed three considerations when he argues that we have no more need of an idea of a perfectly just society to judge locally feasible improvements in justice than we have of a perfectly tall mountain to judge which is the highest actual peak. We will typically need ideal theory to be sure we are “on an acceptable path to a just institutional structure”. Without it, we will struggle to identify the goal towards which we should aim, the “standards of permissibility” governing moves towards that goal and the particular “grievous” injustices whose end we should give priority (Simmons 2010: 34). Without ideal theory, attempts to reform our social and political institutions would lack a target and risk committing wrongs by failing to understand the constraints and priorities they ought to respect. On Simmons’ account, while it might be true that ideal theory is not always necessary for ranking available political reforms, it certainly plays an important role.

This sort of argument seems to me broadly correct (see Jubb 2012). However, it is unlikely that there is any particularly straightforward relation between ideal and non-ideal theory. We should not hope for an algorithm which will convert one into the other, for instance. Consider the complexities of our permissions to use force and deceit against others when the ideal of respecting each other’s rights breaks down. They suggest it will not be easy to derive non-ideal prescriptions from ideal theory, however helpful or even indispensable the latter turns out to be. Tamar Schapiro’s excellent and otherwise very powerful work in this area demonstrates this. It ends up arguing that the letter of an ideal’s rules may have to be violated to somehow respect, preserve or restore their spirit, but is unable to give general rules about how (Schapiro 2003, 2006). The appropriate or permissible responses to the subversion of a practice realizing a given ideal will depend on the ideal, the accompanying practice and how they have been subverted. In that sense, it seems appropriate to focus on the content of potential ideals in order to understand how their spirit might be respected, preserved or restored. One might see, for example, Tommie Shelby’s ‘Justice, Deviance and the Dark Ghetto’ in this light (Shelby 2007).

The second response to critiques of ideal theory that take it to task for not guiding action is more concessive. It accepts that here and now, what we need is a focus on achievable improvements. Instead of exploring the content of unfulfilled ideals to see how to respect, preserve or restore their spirit, this second response mixes ideals with feasibility constraints. For example, for Pablo Gilabert and Holly Lawford-Smith, balancing feasibility against ideal theory makes it possible “to identify the political options that have maximal expected normative



value” and so to “articulate reasonable responses to non-ideal circumstances” (Gilbert and Lawford-Smith 2012: 818, 819). Similarly, for Alan Hamlin and Zofia Stemplowska, providing a theory which gives institutional prescriptions requires marrying feasibility to abstract ideals. An optimal balance between various competing values can be identified once we have drawn the feasibility frontier characterising our situation. Our theory of ideals will provide indifference curves describing the relation between those values. Those will then tell us what the best (set of) point(s) on the feasibility frontier is (Hamlin and Stemplowska 2012: 53ff). That will clear up or at least clarify disputes about ideal and non-ideal theory.

These two articles both describe views which are primarily evaluative and so should be thought of as enactment views. Gilbert and Lawford-Smith straightforwardly assume that evaluative conclusions are at the same time normative ones. It otherwise would not make sense to say that adding feasibility to evaluative rankings of hypothetical worlds makes hypothetical obligations into actual ones without any argument at all (Gilbert and Lawford-Smith 2012: 818-9). If normativity was not entirely derivative of evaluative considerations, then rankings of hypothetical worlds could not by themselves provide hypothetical obligations. Similarly, adding information about which of these worlds could be achieved will not make those obligations actual unless our obligations are defined by evaluations. Gilbert and Lawford-Smith are so committed to the priority of evaluation they obliterate normativity as a distinct category.

Their discussion of Rawls’ choice between different institutional frameworks indicates this is not merely a formal obliteration. Laissez-faire capitalism, state socialism and welfare state capitalism do not fail for Rawls because they do not “provide a maximally desirable feasible instantiation of the demands of political freedom and economic equality” (Gilbert and Lawford-Smith 2012: 820). That makes their failure a matter of not maximizing something and so of degrees. Rawls’ objection to laissez-faire capitalism is not that it gets less equality and liberty than property-owning democracy or liberal socialism. It is that it does not meet the demands of equality and liberty at all. Those demands are specified by the two principles together, and laissez-faire capitalism does not fulfil the two principles (Rawls 2001: 137). Gilbert and Lawford-Smith have imposed an evaluative ranking on a normative requirement.

Hamlin and Stemplowska also operate from a basically evaluative perspective. They describe their model of how ideals interact with feasibility constraints as “teleological and optimising” (2012: 58). Optimising approaches are evaluative, since optimisation requires a ranking and so comparative evaluations. In that sense, Hamlin and Stemplowska are right to worry that their model “may not apply to a deontic approach” (2012: 58). Deontic approaches constrain agents and so are normative, structural views. Their unconvincing response to that worry about the application of their model treats normative requirements as equivalent to external constraints

on ideals. Deontologists do not think of normative requirements as external constraints on ideals though. This is the point of Gilabert and Lawford-Smith's misunderstanding of Rawls. Normative constraints constitute ideals. For Rawls, a society treats its members as free and equal if it is constructed according to the two principles. Normative constraints should not be understood as barriers to achieving ideals in the way that Hamlin and Stemplowska propose.

We can see that feasibility-focused defences of ideal theory like Gilabert and Lawford-Smith's and Hamlin and Stemplowska's are evaluative rather than normative in another way too. These defences are vulnerable to a standard set of objections to maximizing views. An agent maximizing expected justice may well have to violate basic requirements of justice. This is because paths to higher-ranked worlds may well often involve wrongfully removing groups or individuals who obstruct justice. A just world will presumably have fewer violations overall than are required by the path to it, just as the unjust world will anyway contain many more violations. Aiming at justice seems to require the violations. Normative views are not so straightforwardly vulnerable to such an objection. They direct themselves to particular agents and so can require that an agent act in a way that will produce less norm-compliance overall. What other agents do may or may not need to be taken into account in a particular situation by a particular agent. Different effects of acts can be treated differently by a normative view. Rankings of states of affairs seem to have to be impersonal though, and incorporate all features of those states of affairs in the same way, whoever produced them.

Feasibility-focused defences of ideal theory of course want to avoid these objections, but it is not clear that they can. Hamlin and Stemplowska construe normative demands as analogous to feasibility constraints on ideals. That suggests thinking about violating some constraints to secure less violation overall in terms of choosing between being subject to different feasibility frontiers. This makes clear how odd their view is. That choice cannot be made on the basis of the ideals the constraints are supposed to realize. The constraints prevent the realization of ideals rather than being part of them. Does the idea of choosing which feasibility frontier currently applies even make sense? Gilabert and Lawford-Smith instead want to think in terms of Sen's "comprehensive outcomes", which include not just what happens but how it was made to happen (2012: 820). It is not clear how this will help them though. It is not difficult to rank the two relevant comprehensive outcomes. One unjustly perpetuates injustice while the other unjustly brings about justice. If Gilabert and Lawford-Smith's insistence on ranking hypothetical worlds means anything, it must mean that the latter is ranked higher than the former. Alternatively, if the insistence on ranking hypothetical worlds is empty rhetoric, then it is not clear how adding feasibility helps answer the challenge that ideal theory provides no guidance in the here and now.

Both Gilabert and Lawford-Smith's and Hamlin and Stemplowska's models can be fairly straightforwardly understood in terms of David Wiens' "general model" of normative political theory (Wiens 2015: 437). For example, Hamlin and Stemplowska's distinction between the continuum between ideal and non-ideal theory and the separate theory of ideals is similar, if not quite identical, to that which David Wiens draws between directive principles and basic evaluative criteria (Wiens 2015: 435ff). Their disagreement about exactly how to classify Rawls' principle demanding equal basic liberties does not indicate a fundamental disagreement about how to structure a normative theory (Hamlin and Stemplowska 2012: 53; Wiens 2015: 437). The general conception, of empirically-informed directives derived from abstract ideals or values, is clearly shared. Equally, Gilabert and Lawford-Smith's model of rankings of possible worlds supplemented by feasibility constraints to provide directives fits happily with Wiens' account.

Wiens' own commitment to an evaluative and so enactment view is very clear. For him, "a set of directive principles is justified in virtue of the fact that it optimally reflects certain basic evaluative criteria given a set of empirical constraints" (Wiens 2015: 437). The central point of his general model of normative political theory is the justificatory priority of the evaluative over the normative. For Wiens, no norm is justificatorily basic. In this sense, the conclusion Wiens generates from his carefully specified conception of the relation between principles and evaluative judgments is revealing. According to Wiens, the directive principles delivered by ideal theories play no role in providing guidance here and now, and are only likely to mislead. This argument should be particularly troubling for views like Hamlin and Stemplowska's or Gilabert and Lawford-Smith's. It may be that Schapiro, Simmons and others can escape the conclusions of Wiens' argument by rejecting its premises. Hamlin and Stemplowska and Gilabert and Lawford-Smith though seem to share its premises. Their hopes of showing that non-ideal theory is "an extension and complement of, not a substitute for, ideal theorizing" seem doomed by Wiens (Gilabert and Lawford Smith 2012: 819).

### **Not Recipes but Reviews for the Cook-Shops of the Future: Going Evaluative**

Wiens' argument for the conclusion that the directive principles found in theories like Rawls' are useless runs as follows. Any given set of directive principles is justified as a solution to an optimization problem for a particular set of basic evaluative criteria under a given set of empirical constraints. In Rawls' case, the empirical constraints include material scarcity, that those subject to the directive principles have a sense of justice and that the society in which they live is closed (Wiens 2015: 437). Rawls' basic evaluative criteria are freedom and equality. The way the original position models various moral constraints on the choice of principles of justice

supposedly demonstrates this. Those conditions “operationalize” freedom and equality (Wiens 2015: 436). If the argument from the original position works, then its conditions ensure that the choice of principles in the original position optimizes freedom and equality under the relevant empirical constraints. However, for all we know, it only optimizes them under those constraints. Under other constraints, the principles generated and supposedly justified in this way may well not be appropriate. We can only justify using Rawls’ two principles as either a goal at which for reform to aim or a means of ranking available options if we live in a world where Rawls’ assumptions are met. We do not live under those constraints though. Rawls’ principles optimize freedom and equality in conditions which do not apply to us. His theory cannot show that their directives are appropriate for the possibilities created by our different empirical conditions and to which our choices must respond.

As Wiens puts it, directive principles “are “point solutions” – they codify the practical implications of commitment to certain basic values at a particular point within a (constrained) set of possibilities” (2015: 444: italics suppressed). As such, “we cannot justify a reasonable expectation that the normative principles that best reflect our basic values given one set of facts will, given a different set of facts, reflect our basic values to a greater degree than a content-wise dissimilar set of principles” (Wiens 2015: 444). If norms are derivative of values, then norms need to be very carefully designed for the particular circumstances they are expected to govern. If norms need to be very carefully designed for the particular circumstances they are expected to govern, then ideal theories which ignore features of our world cannot play any directive role in it. Responding to critiques of ideal theory by suggesting that adding feasibility constraints will make our ideals action-guiding is misleading.

Wiens’ case for the irrelevance and danger of ideal theory may not be as strong as he claims. At times he seems to stack the deck against his opponents. For example, Wiens’ arguments against what he calls the target and benchmark views seem only to cover at best two of the three roles Simmons gives to ideal theory, that of identifying a goal and perhaps the injustices whose elimination must be made a priority (2015: 440). He does not explicitly address the role ideal theory may have of helping us to understand the constraints on ways of treating each other even in non-ideal circumstances. Both Simmons and Schapiro claim ideal theory has such a role, and seem to have some reason for doing so. The reasoning supporting a prohibition on using physical force to prevent damage to your reputation when principles of justice are perfectly complied with and when they are not will often be similar, for example. Nor does moving from closed to open borders seem to change the way we should resolve conflicts between one person’s bodily integrity and another’s good standing. A prohibition on using aggression towards blackmailers may be a ‘point solution’ to an optimization problem. However, its

rationale seems to partly float free of the conditions characterizing the particular point for which we have justified it.

This point also affects Wiens' arguments against using ideal theory to specify a target at which reform should aim. There, Wiens claims that showing that an ideal theory gives an appropriate target requires showing that its principles characterize the optimal feasible world. Ideal theory is redundant because this requires independently identifying the optimal feasible world (2015: 442). This puts the burden of proof on defenders of ideal theory's guidance role. Wiens effectively asks why we should trust their claims. If there is continuity between the justifications of prescriptions across different worlds, though, we have reason to trust their claims. We will be able to identify appropriate and inappropriate prescriptions because the continuity in the justification will help us identify where a prescription should be applied and where it should not. Wiens' view depends on thinking that we have no reason for thinking a prescription holds outside the strict set of conditions for which it was originally formulated. It would be odd, though, to think that just because a basic structure was not closed, it was no longer a requirement of justice that it not torture its inhabitants, or that slavery might be permissible as long as slaves were born and died elsewhere.

Still, it would be a mistake to dismiss Wiens' argument. He shows that whenever a set of directive principles is justified by a set of evaluative criteria, that justification supposes a set of facts. If those facts do not hold, the justification does not. That may not mean the principles are inappropriate but it should give anyone thinking of using them significant pause. Theorists who share Wiens' commitment to the priority of the evaluative over the normative should pay careful attention to the conditions under and for which they theorise. If they do not, they can have little reason for believing that any prescriptions they make are justified. Instead of optimizing the relevant values for the circumstances in which they intend their principles to apply, they will have created a point solution for another set of conditions.

What about theorists who reject Wiens' commitment to the priority of the evaluative over the normative? Should his arguments persuade them to give up on ideal theory too? Wiens thinks so. He argues that his model is perfectly general, attempting to show that a deontological view like Rawls' can be understood in his terms. If that is right, Simmons and Schapiro do not have the option I mentioned earlier, of rejecting Wiens' conclusion by rejecting one or more of his premises. However, I think Wiens' reconstruction distorts Rawls. A general model should be able to represent any example of what it tries to describe. If Wiens' model seems unable to adequately explain how Rawls' theory works, then it is reasonable to suspend judgment about the generality of his model. If it is reasonable to suspend about the generality of Wiens' model, then his conclusions do not follow for all political theorizing. Theorists committed to the

priority of norms over evaluations would then be justified in continuing to treat ideal theories as capable of appropriately guiding action in a variety of circumstances. Schapiro and Simmons would continue to have the option of rejecting Wiens' conclusions by rejecting one or more of his premises.

I have already noted that Wiens' general account of normative political theory is absolutely clear about its evaluative commitments. Norms are justified only as solutions to the problem of optimizing particular evaluative criteria under given conditions. This does not mean that norms cannot be deontological, as Wiens notes. As he puts it, "directive principles that optimally reflect selected evaluative criteria need not... advise agents to maximize the realization of some moral goal". They may instead "prescribe a set of constraints that prohibits certain kinds of actions whatever the consequences" (Wiens 2015: 439). Still, deontological norms must be justified in virtue of claims about how they "best reflect our moral evaluative criteria" (Wiens 2015: 439). This might be problematic. Moral evaluations are usually impersonal. In contrast, deontological norms constrain the agents they address. That translation may be difficult, since it involves a fundamental shift of perspective. A norm must be required to perform that shift of perspective because evaluations do not direct agents. That norm may of course not be a norm of political or moral theory, but instead of rationality or fundamental agency. Still, Wiens says nothing about why agents must reflect evaluations.

Set aside problems about how norms could be derived from evaluations. The justificatory priority Wiens gives to evaluations seems at odds with the priority of the right commonly associated with deontological and idealizing theories like Rawls'. If "principles of right, and so of justice, put limits on which satisfactions have value" (Rawls 1971: 31), then it is hard to see how those norms could be justified just as reflections of assessments of value. Wiens' evaluative criteria need to be specified independently of the directive principles which supposedly reflect them. Rawls denies this is possible. For example, Wiens' reconstruction of the role of the original position in Rawls' theory treats what Rawls calls 'formal constraints on the concept of the right' as the operationalization of some more basic values. There is no textual evidence for that reading. Rawls thinks of them as general claims about how norms operate, and not as principles operationalizing certain values (1971: 130-136). Wiens' reading of Rawls treats him as fundamentally mistaken about the most basic structure of his theory. Whether the priority of the right is correct, Rawls clearly affirmed it. On Wiens' general model, no political theory affirms the priority of the right. All political theories justify principles of the right as reflections of evaluative criteria.

Wiens' model struggles to interpret other parts of Rawls' theory. Consider how Rawls presents his theory when he describes it as being built around a conception of citizens as free and equal.

That conception demands particular kinds of treatment the more specific principles then describe (see in particular Rawls 2005: 47-88). Freedom and equality here are not functioning as evaluative criteria there. The idea is not that we rank different states of affairs according to how free and equal citizens are in them. That was the mistake Gilabert and Lawford-Smith made when discussing Rawls' views about different social systems. When Rawls talks about citizens being free and equal, he is drawing attention to a status which requires respect. The further principles he advocates make the requirements of that status concrete. That set of norms together realise a more general and abstract norm which ought to govern the relations between a particular set of agents in a particular relationship. The basic feature of Rawls' view is a norm about how agents must treat each other. Since we know that he thought that norms were not justificatorily derivative of evaluations, it seems patronising to insist that he must have been wrong about what he was doing there.

Allowing that norms do not have to be derivative of values can make sense of the role of ideal theory. If the good is not necessarily prior to the right, then norms can be basic. It can be a requirement for particular agents to live under a certain set of norms. That requirement can remain binding even when none of the norms themselves are individually binding. Free and equal citizens must work towards the realization of a basic structure that respects Rawls' two principles because, if Rawls is right, the two principles specify how their political institutions must treat them. Even if Rawls' two principles turn out to be inappropriate for our world, we can learn from them as long as they are appropriate for a world relatively similar to ours. There will be continuity between the rationales for different sets of principles which are appropriate in different circumstances. That continuity gives us a reason to treat principles derived from an ideal theory as relevant. Their rationale will be informative even if they themselves are not appropriate. And anyway, views which treat norms as capable of being justificatorily basic will reject the kind of guidance that makes evaluations primary. In non-ideal conditions, that guidance will tend to call for agents to violate deontological constraints, like Gilabert and Lawford-Smith and Hamlin and Stemplowska do. Theorists who allow that norms can be primary will tend to think that advice is wrong. It does not understand the significance of those deontological constraints.

Yet if Wiens is right about how his supposedly general model functions, this way of working from principles for ideal conditions will not make sense for any theory beginning with evaluations. Ideal theory makes sense as an activity if and seemingly only if you are prepared to accept that norms can be justificatorily basic and so prior to evaluations. Ideal theory is structural, in Bernard Williams' sense. A number of critics of ideal theory appear to reject the idea that norms can be justificatorily basic in that way. Wiens certainly rejects it and there is

strong evidence that Sen, along with Farrelly and perhaps Mills, does too. Their insistence on the priority of evaluations typically plays or seems to play a significant role in their rejection of ideal theory. If we have reasons to accept the possibility of justificatorily basic norms, we also have reasons to reject their attacks on ideal theory. The kind of guidance critics of ideal theory want will seem instrumental and inappropriate if norms can be basic. Equally, it will seem possible to give advice which does respect norms on the basis of ideal theorizing.

I have not tried to show that norms can be justificatorily basic here. None of the pieces I discuss here argue the contrary though. Wiens tries to show that his model is general by using it to reconstruct Rawls' theory. I have argued that reconstruction is distorting. Even if it was accurate, that would not be an argument that norms cannot be justificatorily basic. It would at best show that treating norms as always justified by evaluations does not prevent you making sense of Rawls' view and others structurally similar to it. Nor is it clear how an argument that norms cannot be justificatorily basic would run. For one thing, it would have confront the problem that arguments are supposed to compel agents' beliefs and so presuppose epistemic or justificatory norms. In that sense, the discussion here has reached a kind of impasse. The disagreement seems to be the result of disagreement about axiomatic or foundational commitments.

Nonetheless, in what remains I hope to suggest that part of the case for the superiority of treating evaluations as primary is not as strong as its advocates suppose. To do so, I evaluate that case by the criterion it uses for its opponents. Ideal theory is supposedly inadequate because it does not generate action-guiding prescriptions. I have suggested that depends on thinking that prescriptions must have a certain character, a character it is not clear they must have. Nonetheless, ideal theory is not particularly successful at generating a particular kind of action-guiding prescriptions in non-ideal circumstances. Work by advocates of ideal theory like Schapiro shows that. There is not an algorithm to derive a ranking of alternative policies from an ideal theory of justice. The critique of ideal theory by those who, like Sen and Wiens, seem to presuppose its failure could still stand. Even if ideal theory is not as bad at guiding action here and now, it might still be worse than the primarily evaluative theory Sen and Wiens urge it should be replaced with. We should check that accepting a primarily evaluative perspective would in fact generate the kind of action-guiding prescriptions critics of ideal theory demand. Are they able to meet the challenges they raise for it? Can their arguments avoid defeating themselves? There seem to be two reasons for scepticism.

The first problem for evaluative critiques of ideal theory is that assessments of the possibilities open to us are very difficult to make accurately. We tend to systematically fail to understand the set of effects our acts will bring about or the way that we will respond to those effects. If we are



to evaluate options though, we need to know what their effects will be. We cannot assign evaluative significance to costs and benefits we cannot even identify. Without accurate predictions, rankings of alternatives can hardly be complete. Nor is this such a serious problem for transcendental ideal theory, which limits the set of effects it is concerned with in two ways. First, once we can be sure that an option realizes a transcendental ideal, a theory calling for that realization needs no further information. It does not need to know as much about alternatives as an evaluative theory which needs to take into account any incremental improvement in a set of values that may range over a very wide variety of effects. Second, ideals direct particular agents to act in particular ways. Agents must relate to each other in certain ways. The content of these acts may only depend on what other agents do in a very restricted way, if at all. Norms forbidding intentional killing except in a very restricted set of circumstances do that and only that. The range of effects an act must take into account to comply with a norm can be very limited. This is the point of the attraction of quasi-deontological rules of thumb for consequentialists. Difficulties calculating effects do not have to be as significant, because fewer effects have to be calculated. The shadow of the future can be much less consequential.

The second difficulty faced by a project of guiding action here and now by ranking the available options is of identifying and applying the relevant evaluative criteria. Wiens' model treats directive norms as justified as optimizations of more basic values. All primarily evaluative theorists see that relation or one like it as justifying norms. That is what makes someone a primarily evaluative theorist. However, that relation depends on a set of empirical conditions in which the norms in fact do optimize or otherwise reflect the more basic values. Moving appropriately from norms to evaluative criteria means identifying what those conditions are and how the norms respond to them. A concrete normative judgment about a particular situation can only be used to understand the relevant evaluative criteria if we know under exactly which conditions it applies. Wiens' argument about the uselessness of ideal theory depends on that claim. Otherwise ideal theorists would be entitled to assume that their principles appropriately reflect basic evaluative criteria under a variety of conditions. And if they could assume that, then ideal theory would not be useless. It would not be tied to a particular set of conditions.

This makes understanding our values potentially very difficult. In order to understand our view that slavery is wrong, for example, we need to know whether that is a judgment about slavery in general, or a particular set of exemplars like, for instance, that found in the southern United States of America before the Civil War. Otherwise our inferences about the shape and content of the value or values it relates to will be wrong. Even once we have identified in which circumstances a judgment applies, we still need to know how it responds to them. And

separating these two tasks out will not be straightforward at all. The circumstances in which a norm responds to a value will depend on how we think it responds. Our most basic moral and political commitments, like Rawls' view, taken from Lincoln, that "if slavery is not wrong, nothing is", can no longer serve as straightforward evidence for what we ought to do (Rawls 2001: 29). Instead of serving as *prima facie* identifications of constraints that must be observed, their role is severely circumscribed. Those commitments must be connected them to basic values by identifying the conditions under which they optimize those most basic values. At the same time, we need to understand how they reflect those values. Identifying the shape of those basic values may then be very difficult.

The same operation needs to be performed in reverse in order to apply those values and so produce a ranking. The relevant features of a state of affairs need to be assessed in terms of our basic evaluative criteria. I have already suggested that this is likely to be very difficult because identifying what has to be ranked is likely to be very difficult. However, it will also face the problem of what applying our basic values to an identified set of acts, events and effects actually means. Sen's examples are misleading here. We do not need theorising to tell us that, all other things being equal, a reduction in illiteracy or hunger is an advance. That is like the unacceptability of slavery: it is the kind of thing which we know already, if we know anything at all about how our polities ought to be organized. We would reject theories if they implied that, all other things being equal, a reduction in illiteracy or hunger was not an advance. We need theorising to help us deal with more complex questions. An evaluative theorist needs to know what the relative importance of a reduction in illiteracy or hunger is when it comes at the cost of holding back some other goal or ideal. Even deciding how much of a seemingly straightforward value like welfare a given state of affairs realizes may be incredibly difficult and highly contentious. More complex candidate values like freedom or equality will only going to generate even more serious problems. There is no obvious reason to suppose that applying these criteria to complex political questions to evaluate alternatives is going to be possible at all, even if we can identify all the relevant features of those alternatives.

If ideal theories are useless because they do not tell us how to rank available policy alternatives here and now, then it is not clear how moving to a primarily evaluative view helps. When we need advice at all, it seems such a view needs both more empirical and conceptual or evaluative information than we are likely to be able to supply. Advocates of evaluating feasible social worlds criticize ideal theory because it does not do so. However, even accepting that norms must respond to evaluations does not mean that seeking such evaluations will provide us with helpful prescriptions about how to improve our social world. Views which treat norms as basic may well be no worse off. There are two serious problems subordinating norms to values

creates. First, a wider range of effects need to be taken into account by evaluative views, and second, identifying basic values and the norms that realize them may be very difficult. If writing recipes for the cook-shops of the future is superfluous to actually trying to create them, it is hardly clear why reviewing them is a sensible alternative instead.

## **Conclusion**

I have been mainly concerned to try to show two things here. First, I wanted to show that a primarily evaluative position struggles to make sense of one understanding of ideal theory. If ideal theory is about perfection, then an evaluative stance, which expects to rank states of affairs and their contents, will make it seem pointless. Ideal theory in Sen's sense of transcendental theory answers a question which is only really comprehensible from a position which denies that the good is necessarily prior to the right. Second, I wanted to suggest that primarily evaluative positions have various problems, similar to those with which critiques of ideal theory tasked its adherents. It may well be very difficult to give concrete prescriptions about what to do in a non-ideal world if you begin by thinking that norms are derivative of evaluations. You will need information that may be quite beyond our capacities to acquire, both about features of states of affairs and of our basic evaluative criteria.

My aim in trying to show these two things is to expose what seem to me a number of misconceptions in the literature on the ideal – non-ideal theory controversy. If I am right about the first of my two claims, it is understandable that theorists like Sen and Wiens reject ideal theory. Their position on a seemingly separate question determines that rejection. Theorists who deny that evaluations are justificatorily prior can legitimately mobilise a set of arguments that may, on the face of it, have little to do with the ideal – non-ideal controversy to defend their position in that debate. Such theorists may also worry that ideal theory cannot provide concrete guidance of the sort primarily evaluative theorists demand. They should reassure themselves that primarily evaluative theorists are likely to be worse off. They can then move to the task Schapiro identified more than a decade ago, of thinking through the difference between the spirit and the letter of ideals. In particular, theorists who allow that norms may be justificatorily basic should focus on this instead of heading up what seems to me for them the blind alley of primarily evaluative discussions of feasibility.

## Bibliography

Colin Farrelly (2007). 'Justice in Ideal Theory: A Refutation', *Political Studies* 55 (4):844-864.

William Galston (2010). 'Realism in Political Theory', *European Journal of Political Theory* 9 (4):385-411.

Pablo Gilabert and Holly Lawford-Smith (2012). 'Political Feasibility: A Conceptual Exploration', *Political Studies* 60 (4):809-825.

Alan Hamlin and Zofia Stemplowska (2012). 'Theory, Ideal Theory and the Theory of Ideals', *Political Studies Review* 10 (1):48-62.

Robert Jubb (2012). 'Tragedies of Non-Ideal Theory', *European Journal of Political Theory* 11 (3):229-246.

Robert Jubb (2015a) 'The Real Value of Equality', *Journal of Politics* 77 (3):679-691.

Robert Jubb (2015b) 'Playing Kant at the Court of King Arthur', *Political Studies* 63 (4): 919-934.

Charles Mills (2005). "'Ideal Theory" as Ideology', *Hypatia* 20 (3):165-184.

John Rawls (1971). *A Theory of Justice*, Cambridge, Mass.: Harvard University Press.

John Rawls (2001). *Justice as Fairness: A Restatement*, Cambridge, Mass.: Harvard University Press.

John Rawls (2005). *Political Liberalism*. New York: Columbia University Press.

Enzo Rossi and Matt Sleat (2014). 'Realism in Normative Political Theory', *Philosophy Compass* 9 (10):689-701.

Tamar Schapiro (2003). 'Compliance, Complicity, and the Nature of Nonideal Conditions', *Journal of Philosophy* 100 (7):329-355.

Tamar Schapiro (2006). 'Kantian Rigorism and Mitigating Circumstances', *Ethics* 117 (1):32-57.

Amartya Sen (2006). 'What Do We Want from A Theory of Justice', *Journal of Philosophy* 103 (5):215-238.

Tommie Shelby (2007). 'Justice, Deviance and the Dark Ghetto', *Philosophy and Public Affairs* 35 (2):126-160.

A. John Simmons (2010). 'Ideal and Nonideal Theory', *Philosophy and Public Affairs* 38 (1):5-35.

Laura Valentini (2012). 'Ideal vs. Non-Ideal Theory: A Conceptual Map', *Philosophy Compass* 7 (9):654-664.

David Wiens (2015). 'Against Ideal Guidance', *Journal of Politics* 77 (2): 433-446.

Bernard Williams (2005). 'Realism and moralism in political theory', in Williams, ed. G. Hawthorne, *In the Beginning Was the Deed*. Princeton University Press, pp. 1-17.