# Will The Real Principles of Justice Please Stand Up?

David Wiens

**Abstract.** This chapter develops a "nesting" model of deontic normative principles (i.e., principles that specify moral constraints upon action) as a means to understanding the notion of a "fundamental normative principle". I show that an apparently promising attempt to make sense of this notion such that the "real" or "fundamental" demands of justice upon action are not constrained by social facts is either self-defeating or relatively unappealing. We should treat fundamental normative principles not as specifying fundamental constraints upon action, but as specifying basic criteria for comparatively evaluating and ranking possibilities.

Philosophers widely concede that what an agent ought to do depends on the options that are, in some sense, available to her. Even full-throated defenses of so-called ideal theory concede that our theorizing about what we ought to do "here and now" — as a matter of policy, in the face of manifestly unjust and otherwise nonideal circumstances — must be sensitive to myriad social facts. Yet, they contend, the "real" or "fundamental" demands of justice do not concede to such facts so readily (e.g., Cohen, 2008; Estlund, 2014). Justice stands among the highest virtues; her clothes must remain unstained.

The notion of "fundamental demands of justice" or "fundamental normative principles" underlying this Nonconcessive Refrain is a vague and, I think, elusive one. In this chapter, I present a general framework for making sense of this idea. In so doing, I show that an apparently promising attempt to elucidate the Nonconcessive Refrain in terms of deontic normative principles — i.e., principles specifying constraints upon action — is either self-defeating or relatively unappealing. We can do better, I will suggest, if we treat fundamental normative principles as specifying basic normative criteria for comparatively evaluating and ranking possibilities.

Here's a high-altitude map of the road ahead. I start by distinguishing between two kinds of normative principle: one kind performs an evaluative function, the other performs a deontic (or, as I will say, "directive") function. I then articulate what I will call the Uncontroversial Thesis: that directive principles are specified relative to a particular set of salient possibilities (leaving salience vague for now). This is followed by a brief summary of an "optimization model" of normative theories, the details of which are presented at greater length elsewhere (Wiens, 2015*a*). My sole aim here is to show that the

model offers a compelling general explanation of the Uncontroversial Thesis while also making the implications of that thesis precise. Following a summary of the optimization model, I take up the chapter's key issue: namely, how we might make sense of the vague notion of a "real" or "fundamental" moral directive. To this end, I present a "nesting" model of directive principles, whereby one set of directive principles $P$ is subordinate to another $P'$ if the set of possibilities relative to which $P$ is specified is a proper subset of the set of possibilities relative to which $P'$ is specified. On this model, fundamental directives are identified with maximally superordinate directives; that is, fundamental directives are those that are specified relative to a maximally encompassing superset of possibilities. At last, I consider a promising way to make sense of the notion of a "maximally superordinate directive" and show that this is either self-defeating or leaves us with a relatively unappealing account of fundamental directive principles. I close by suggesting that a more plausible account of fundamental normative principles treats these as evaluative rather than directive principles.

Disclaimer: This chapter is a companion to Wiens (2015*a*), in that the arguments to follow draw heavily on the details of the optimization model therein, while also exposing further implications of that model. Space limitations require that I merely summarize the key ideas of the optimization model here, at the risk of demanding too much of the reader's attention. That model nonetheless derives support from the fact that it offers a compelling explanation of the Uncontroversial Thesis.

## 1. TWO KINDS OF NORMATIVE PRINCIPLE

Let's start by distinguishing between two kinds of normative principles according to their function in normative theorizing.

*Evaluative principles* serve to comparatively assess and rank options according to some set of normative criteria, where the relevant options include actions, institutional schemes, or states of affairs. The relevant set of normative criteria comprises basic moral and social values, like liberty, equality, welfare, and so on. Stated thus, basic evaluative criteria are too general to yield anything more than the crudest of evaluative judgments — for example, that a democratic society is more equal (in some sense) and exhibits greater respect for individual freedom (in some sense) than a slave-holding society. Moreover, there is much dispute about how to best conceptualize and operationalize basic values like liberty, equality, and welfare. Take equality, for instance. Luck egalitarians contend that it is best understood as a property of social distributions, namely, that distributive inequalities are justifiable only if they are the consequence of voluntary choices (see, e.g., Cohen, 1989). Relational egalitarians, in contrast, argue that equality is best understood

as a property of social relations, namely, that fellows in a cooperative scheme not be subject to oppressive or demeaning social hierarchies (see, e.g., Anderson, 1999). At a minimum, evaluative principles specify a candidate interpretation of the basic criteria we use to comparatively assess options. They must also specify the relative significance of basic values: How does equality relate to other values, like liberty and welfare? Are some basic values more important than (i.e., weightier) than others? If so, which ones? What's the relative weight of these more important ones? If not, is there simply a plurality of equally important basic values? How, if at all, do we aggregate across these? These sorts of questions are settled by evaluative principles.

In contrast with the ranking of options delivered by evaluative principles, *directive principles* mark the lines between obligatory, permissible, and impermissible options. Directive principles perform a deontic function: they pick out a subset of options and identify these as the options that, from a normative standpoint, are to be realized. Directive principles thereby specify constraints on permissible conduct. For instance, if we take Rawls's two principles of justice to be correct, and if we take justice to be a deontic concept, then we arrive at a normative injunction to coordinate our individual behavior so as to realize (perhaps only eventually) an institutional scheme that satisfies Rawls's two principles. This is not simply an evaluative judgment — namely, that Rawls's two principles characterize the most morally desirable institutional scheme. This is a deontic judgment — one acts permissibly only insofar as one's conduct is consistent with a requirement to coordinate on an institutional scheme that satisfies Rawls's principles; otherwise, one acts impermissibly.[1]

A general principle is not identified as evaluative or directive by virtue of its content but by virtue of the function it serves within a normative theory. A principle of strict distributive equality can serve an evaluative purpose without entailing anything about permissible action: e.g., judging that a state of affairs $s$ is normatively superior to an alternative $s'$ insofar as $s$ satisfies a principle of strict equality to a greater degree than does $s'$ need not entail that the only permissible states of affairs, from the standpoint of justice, are those that satisfy a principle of strict equality. Alternatively, a principle of strict equality can serve a deontic purpose without entailing anything about comparative evaluations: e.g., judging that a state of affairs $s$ is permissible only if $s$ satisfies a principle of strict equality need not entail that $s$ is higher ranked, from the standpoint of justice, than an alternative $s'$ that satisfies a principle of strict equality to a lesser degree. There's a

---

1   There is surely an implicit domain restriction here that provides for a minimally adequate sphere of private conduct: roughly, one acts (im)permissibly within the "public sphere", the domain of conduct that is governed by a requirement to coordinate on a particular institutional scheme.

simple reason why evaluative judgments might not closely track with deontic judgments: the latter are subject to considerations to which the former is not. For instance, as seems reasonable to many: which options ought to be realized are subject to feasibility considerations in some way, while a ranking of options need not be sensitive to such constraints.

## 2. THE UNCONTROVERSIAL THESIS

The starting point for this chapter is what I take to be the following Uncontroversial Thesis: *a set of directive principles is justified relative to a particular set of salient possibilities.* Let's refine that: which options — actions, institutional schemes, states of affairs — an agent is required, permitted, or prohibited to realize depends on the set of options that is open to that agent in some (for now, unspecified) sense. This thesis emerges from numerous quarters in moral and political philosophy.

Consider the dictum that "ought implies can": what one ought to do depends on what one can (in some sense) do (e.g., Vranas 2007; but see Graham 2011).[2] Or consider cases like Professor Procrastinate (Jackson and Pargetter, 1986). By assumption, Procrastinate ought to accept the review assignment relative to the set of possible worlds at which he will complete the review; yet many agree that he ought not accept the assignment relative to the set of worlds at which he will fail to complete the review. So the directive principle to which Procrastinate is subject is sensitive to the set of possibilities one deems salient.[3] In a similar vein, consider the case of the ten miners who could be trapped in one of two shafts, *A* or *B* (see Kolodny and MacFarlane, 2010). If we block the shaft in which the miners are located to protect them from an oncoming flood, all ten are saved; if we block the wrong shaft, all ten die. So if we know which shaft they occupy, we ought to block that shaft. Yet, plausibly, we ought not block either shaft when we are uncertain about the location of the miners; in which case only one of the ten will die. (The relevant sets of possible worlds being circumscribed here by our evidence about the location of the miners). We can proliferate similar cases to further illustrate the point (see, e.g., Goldman, 1976).

Turning elsewhere, Rawls's restrictions on the application of his two principles of justice nicely expresses the Uncontroversial Thesis:

---

2  "Ought" is typically taken to be a deontic category, although it can be understood evaluatively too (see, e.g., Cohen, 2008, chs. 6, 8). To avoid confusion, I use "ought" to denote a directive principle throughout.

3  It's a distinct matter whether the requirement to accept *and complete* the review obtains at the worlds at which Procrastinate fails to complete the review (see Estlund, 2011, 216ff). I set this aside until later.

> By putting these principles in lexical order, the parties are choosing a conception of justice *suitable for favorable conditions.* ... [W]e must still ask how well they apply to institutions under less than favorable conditions, and whether they provide any guidance for instances of injustice. *The principles of justice and their lexical order were not acknowledged with these situations in mind and so it is possible that they no longer hold.* (Rawls, 1999, 215–216, emphasis added)

That is, if we specify directive principles of justice under assumptions that differ from Rawls's own assumptions — under circumstances that imply a set of possibilities that are less hospitable to the realization of, e.g., equal basic liberties or a sufficient level of economic development — we might well wind up endorsing different directive principles.

A similar thought underlies much recent discussion of the assumption of highly idealized social conditions in so-called "ideal theory". Valentini's distinction between "good" and "bad" idealizations in defense of the practical relevance of ideal theory is representative here (Valentini, 2009, secs. 4–6). Good idealizations merely simplify reality by abstracting from certain complicating features; bad idealizations, in contrast, "distort" reality at a basic level. Good idealizations leave us with a model world that, despite its simplicity, is a credible model of *our world* and can thus accommodate the later introduction of real world complications without fundamentally altering the associated normative theory. Badly idealized model worlds are not credible models of our world but of fundamentally distinct *alternate realities*; such models can accommodate later introduction of real world complications only at the cost of fundamentally altering the associated normative theory. I don't wish to interrogate Valentini's distinction here (instead see Jacopo Uberti, 2014). The relevant point is that much of the hand-wringing regarding the use of idealizations in normative theory acknowledges (if only implicitly) that a normative theorist's assumptions about the set of feasible options — as implied by their assumptions about the causal mechanisms that produce social outcomes — are consequential for the content of the resultant normative theory. In short, the worry is that different assumptions about what's possible lead to different normative directives.

Finally, notice that the standard semantics for deontic modals (e.g., "ought", "may", etc.) aligns with the substantive thought proposed here. On the standard view, "*A* ought to $\phi$" is analyzed thus: *A* performs $\phi$ at all of the highest-ranked worlds within a set of salient worlds (see, e.g., Kratzer 1991; Charlow forthcoming provides helpful discussion). Which worlds are ranked highest (as determined by the relevant evaluative criteria; "ordering source" on the standard terminology) depends on the set of worlds that is picked out as "salient" (those that are consistent with the relevant facts or the available evidence or

whatever; the "modal base" on the standard terminology). So the standard semantics implies that the *semantic* content of the terms we use to express obligations, permissions, and so on — as conveyed by directive principles — is sensitive to the set of possibilities taken to be, in some sense, available. It would be surprising to find that the *substantive* content of our directive judgments is not similarly sensitive to the set of possibilities taken to be, in some sense, feasible.

## 3. AN OPTIMIZATION MODEL OF NORMATIVE THEORIES

Summarizing thus far: The substantive content of directive principles is sensitive to some salient set of possible options, those that are open or available (in some sense) to the relevant agents. This is the Uncontroversial Thesis. A compelling explanation for this thesis can be found in the "optimization model" of normative theories I present elsewhere (Wiens, 2015*a*). I do not argue here that the optimization model is *the correct* way to think about the structure of normative theories. I take the model as given and simply summarize its key points. The model finds support here insofar as it offers a compelling explanation for a thesis that finds support across many quarters of normative philosophy. Ultimately, my objective is to articulate the Uncontroversial Thesis precisely and show that a compelling explanation for that thesis leads to a relatively unappealing account of fundamental directive principles.

The basic thought is that we can understand the Uncontroversial Thesis by viewing normative theories as a certain kind of model. In brief, normative theories conjecture normatively optimal solutions to well-specified problems. To illustrate the point, consider two familiar examples. On my proposal, Rawls's two principles conjecture a normatively optimal solution to the problem modeled by the original position: roughly, the specification of terms of social cooperation that free and equal persons can readily accept in view of certain "circumstances of justice". Similarly, Nozick's theory of the libertarian minimal state conjectures a normatively optimal solution to the problem of biased and unreliable enforcement of people's natural rights within a Lockean state of nature. In each of these examples, a normative theory conjectures a set of directive principles as the best way (from a normative standpoint) to solve the specified problem. (They also propose a particular problem as *the* salient problem from the standpoint of justice, but I set this aside here.)

Hopefully these examples provide a clear sense of what it would mean to treat a normative theory as conjecturing a solution to a specific moral problem (see Wiens 2015*a* for a more detailed discussion). But how are we to understand the idea of a *normatively optimal* solution? Recall the difference between evaluative principles and

directive principles. Evaluative principles rank options in accordance with a specified set of evaluative criteria (e.g., liberty, equality, security, and so on); while directive principles demarcate the subsets of options we are required (permitted, prohibited) to realize. The notion of a "normatively optimal solution" is understood in terms of the ranking implied by evaluative principles: normative theories conjecture a set of directive principles as the solution to a specific problem — in particular, a set of directive principles that ranks highest among the available options in light of certain evaluative criteria. Put differently, normative theories proffer a set of directive principles that, if satisfied, would bring about a state of affairs that, among the available alternatives, realizes the specified evaluative criteria to the highest degree.

We can make this point precise by appeal to an "optimization model" of normative theories. On this model, normative theories have two inputs: a set of evaluative principles and a set of feasibility constraints. We have already discussed the key features of evaluative principles above. The important point here is this: we say that an option (action, institutional scheme, state of affairs, etc.) is normatively superior than an alternative if and only if the specified evaluative principles imply that the former is ranked above the latter.

While evaluative principles serve to rank options, feasibility constraints serve to exclude options from consideration. How, precisely, we analyze the relevant notion of feasibility doesn't matter for my purposes here.[4] All that matters is that an option is considered feasible only if it is consistent with certain specified constraints on the realization of basic evaluative criteria and that infeasible options are considered "unavailable" for the purposes of a particular theoretical task. For instance, Rawls's assumptions that society is closed to cross-border transactions excludes states of affairs involving cross-border transactions from consideration in the original position. For the purposes of identifying principles of distributive justice for a society, such states of affairs are considered "unavailable" (or: their availability is considered irrelevant). As a result, comparative assessment of alternative sets of principles of justice in the original position is limited to considering the consequences of satisfying those principles at states of affairs in which interpersonal transactions are limited by state borders.

Whereas evaluative principles and feasibility constraints are the inputs for normative theorizing, a set of directive principles is the output. Recall that the key function of directive principles is to demarcate the subsets of options that are required, permissible, and prohibited. On the optimization model, such principles are specified in light of the

---

4  For some recent discussion on the topic, see  Gheaus (2013); Gilabert and Lawford-Smith (2012); Lawford-Smith (2013); Wiens (2015*b*).

ranking implied by the specified evaluative principles and the limits set by the specified feasibility constraints. In brief, a set of directive principles is specified by reference to the set of optimal feasible options, where "optimal" is understood in terms of the specified evaluative principles. There are several ways to elaborate this idea. Here are a few: A set of directive principles is to be endorsed as specifying normatively appropriate constraints on action in virtue of the fact that it characterizes, in a general way, the core institutional and behavioral features of the optimal feasible states of affairs. Alternatively: a set of directive principles is justified in virtue of the fact that it best characterizes the deontic implications of a commitment to certain basic values given certain limitations on the realization of those values. Or, finally: a set of directive principles is justified in virtue of the fact that it conjectures a solution to a specific problem that best realizes certain evaluative criteria under the specified feasibility constraints. These are all mutually consistent expressions of the proposed model's core idea.[5]

The point to emphasize here is that the optimization model elucidates and explains the Uncontroversial Thesis insofar as it implies a plausible articulation of that thesis: according to the model, the specification of a set of directive principles is sensitive to the specification of feasibility constraints. This follows straightforwardly from three premises:

(1) Certain facts — e.g., those pertaining to the availability of material resources, or people's propensities to respond to different kinds of incentives, or the set of candidates for cooperative partnership — constrain the possibilities for realizing the specified evaluative criteria.

(2) Directive principles are specified in view of a comparative assessment of states of affairs that satisfy a specification of these feasibility constraints — they are conjectures about the core institutional and behavioral features of the feasible states of affairs that best realize certain specified evaluative criteria.

(3) The institutional and behavioral features of the states of affairs that best realize the specified evaluative principles are liable to change in response to changes in the set of possibilities for realizing the specified evaluative criteria.

The first premise is just the (uncontroversial) claim that attempts to realize certain moral and social values at any world are subject to feasibility constraints. The second premise is a statement of the central idea of the optimization model.

---

5   An important concern that might arise here is that the model's optimizing structure biases it in favor of consequentialist normative theories or represents an attempt to "consequentialize" nonconsequentialist theories (on "consequentializing", see Portmore, 2009). See Wiens (2015*a*, 439) for my response.

The third premise is the claim that *how* the optimal balance of basic evaluative criteria is realized at any world depends in some respect on how that world lies. I think this claim, too, is uncontroversial. Consider the following illustration. Suppose we have three evaluative principles: a state of affairs $s$ is normatively superior to an alternative $s'$ insofar as (1) the distribution of the relevant goods is more equal in $s$ than in $s'$; (2) the level of total welfare realized in $s$ is higher than in $s'$; and (3) the degree to which people's labor/leisure choices are subject to state coercion in $s$ is less than in $s'$. (Assume further that none of these principles are given lexical priority; thus, some as-yet unspecified tradeoffs are permitted at the margins.) If we confine our attention to the set of worlds (denoted $W$) at which people's labor productivity is insensitive to the *relative* size of their reward packages, then an institutional scheme that leaves everyone with strictly equal post-tax income will maximally realize the basic value of strict distributive equality without sacrificing the level of total welfare realized and without requiring the state to coerce people to make productivity-maximizing choices. For this set of worlds, the optimal realization of the specified evaluative criteria invariably involves implementing a taxation scheme that maximally realizes post-tax distributive equality (subject to technological limitations). However, for a set of worlds at which labor productivity is incentivized by relative gains (denoted $W'$), an institutional scheme that equalizes post-tax income realizes distributive equality at the expense of total welfare or by subjecting people's labor/leisure choices to a notable level of state coercion. For the set of worlds involving these tradeoffs, realizing the optimal *balance* of values will likely require an institutional scheme that sacrifices along all three dimensions. Certainly, it will not involve a taxation scheme that maximally realizes post-tax distributive equality. In general, then, which institutional mechanisms engender the optimal balance of evaluative criteria depends on the background conditions in which those institutions are expected to operate.

What follows from the preceding is a precise articulation of the Uncontroversial Thesis: the specification of directive principles is sensitive to the specification of a salient set of possibilities, which is circumscribed by a specification of feasibility constraints. Continuing with the preceding example: Suppose $W$ and $W'$ are included in the set of feasible worlds (i.e., the worlds in both $W$ and $W'$ are consistent with a theorist's assumptions about the relevant empirical matters). As the worlds in $W$ are ranked higher than those in $W'$ according to the specified evaluative principles, the resultant set of directive principles will be specified by appeal to the core institutional and behavioral features of the worlds in $W$. Without an in-depth analysis of the worlds in $W$ (which is beyond the scope of this paper), we can only speculate about their core features; but, plausibly, they contain societies that realize socialist aspirations to a high degree  (cf.

Cohen, 2009). Thus, plausibly, we arrive at socialist directive principles in this case (e.g., Carens, 1981). Now suppose that the worlds in $W$ are excluded from the feasible set (i.e., they are inconsistent with a theorist's assumptions about the relevant empirical matters) and assume that some subset of $W'$ composes the set of morally optimal feasible worlds. In this case, the resultant directive principles will be specified by appeal to the core institutional and behavioral features of the worlds in the optimal subset of $W'$. Again, without an in-depth analysis of the optimal worlds in $W'$, we can only speculate about their core features. Moreover, our ranking of worlds within $W'$ will depend on the relative weight assigned to the specified evaluative principles, so that we can comparatively assess the different ways of trading off among the three criteria. In any case, it seems implausible that the optimal $W'$ worlds will approximate socialist utopias; thus, it seems implausible that we will arrive at socialist directive principles in this case.[6]

In general: Given the optimization model of normative theories, we can understand the Uncontroversial Thesis as asserting that a set of directive principles is justified — i.e., is to be endorsed for the purposes of assigning obligations, permissions, and so on — relative to the set of possibilities that is consistent with certain specified feasibility constraints. As shorthand, we say that *directive principles are constraint-relative.*

## 4. A NESTING MODEL OF DIRECTIVE PRINCIPLES

Philosophers tend not to baulk when they encounter the claim that "oughts" are constraint-sensitive in the above-modeled sense. Yet there are cases (e.g., Professor Procrastinate, the Miner case) in which agents appear subject to competing directives. Philosophers often deal with these conflicts by introducing distinctions among "oughts", like "all-things-considered oughts" versus "moral oughts", or "nonideal oughts" versus "ideal oughts", or "concessive oughts" versus "nonconcessive oughts". Consider a typical response to Professor Procrastinate:

> Sure, given that he won't complete the review, he ought to refuse. But

---

6   Perhaps the optimal $W'$ worlds are Rawlsian well-ordered societies, in which case, we arrive at Rawlsian directive principles; perhaps they are right-libertarian utopias, in which case, we arrive at right-libertarian directive principles. The results will be shaped by the evaluative principles that guide reflection on the required tradeoffs, which are left unspecified here. The key point pertaining to feasibility constraints is that different assumptions about feasibility will impose different kinds of tradeoffs. Insofar as the specification of directive principles proceeds in light of a comparative assessment of different ways of trading off evaluative criteria, the fact that different specifications of the feasible set leaves us with different options for trading off evaluative criteria implies that the specification of directive principles is sensitive to the specification of feasibility constraints (cf. Cowen, 2007). This is just a different way of stating the main point of this section.

that's just a statement about what he ought to do, all things considered, in light of certain nonideal facts about his character. An obligation to refuse is a concession to Procrastinate's avoidable moral deficiencies. Advising that he refuse the assignment might make for good policy, but it can hardly be the case that *morality* lets him off the hook so easily. His *moral* obligation is to accept the review and complete it.

Here's the underlying thought, stated more generally: When considering how to best proceed in a particular situation, we should adopt the course of action that yields the best expected outcome from the perspective of morality.[7] This requires attending to the probability distributions over the outcomes that might arise from competing courses of action, which requires attending to myriad facts about the world. This is simply good policy: when deciding what to do, we should attend to how the world lies, so to speak. But reflection on the *fundamental* demands of morality or justice should not concede so readily to the facts—especially facts about our moral deficiencies.

Many people find this thought quite reasonable, its vagueness notwithstanding. One potentially helpful way to represent the requisite notion of a fundamental obligation uses the picture of a "hierarchy of requirements" (Estlund, forthcoming). On this picture, requirements (i.e., directives) are arranged according to their "degree of concession" or the degree to which the constraints under which they are specified reflect nonideal features of the world. More concessive or nonideal directives are then said to be subordinate to less concessive (more ideal) directives. Given such a hierarchy, we might naturally think that the "real" or "fundamental" directives of morality (or justice or whatever) are the maximally superordinate ones ("maximal directives" for short).[8]

Reflection on cases like Professor Procrastinate might convince us that directive principles bear this hierarchical structure. Procrastinate has a concessive obligation to refuse the assignment in light of his anticipated failure to deliver; yet there stands a superordinate nonconcessive obligation to accept and complete the assignment. As Estlund puts it (following Jackson and Pargetter (1986)), the nonconcessive obligation

---

7 This need not be read in a consequentialist manner if we understand outcomes "comprehensively", that is, to include facts about *how* a certain state of affairs was brought about. See Sen (2002, chs 1 and 4) on comprehensive outcomes.

8 One might be tempted to think that the *sufficiently* superordinate ones are good enough. Suppose the hierarchy has $n$ levels, with $n$ representing the maximally superordinate level. Suppose we say that directives at $n-3$ are sufficiently superordinate to qualify as the "real" moral directives. Why $n-3$? Why not $n-4$ or $n-2$? And, if we fix $n-3$ as the sufficiency threshold, are the directives at level $n-2$ "more fundamental" than those at $n-3$ in virtue of being superordinate? Unless there is a principled way to fix the threshold at some level $m < n$, the fundamental moral directives must be identified with the maximally superordinate ones.

has "primacy" because, at a world where Procrastinate completes the assignment, the concessive obligation to refuse "evaporates", it fails to obtain. In contrast, at the nonideal world where Procrastinate fails to complete the assignment, the nonconcessive obligation to accept and complete the assignment continues to obtain (this volume, XX).

As things stand, though, we are merely asserting a hierarchy of principles that grants primacy to nonconcessive obligations on the basis of one especially suggestive example. Can we make clear sense of the relevant notion of primacy more generally? To do so, I propose a "nesting" model of directive principles, whereby directives appear hierarchically-arranged with respect to each other because they are specified relative to nested sets of possibilities. Continuing with the Procrastinate example, let $C$ be the set of whatever facts about Procrastinate determine that he will not complete this particular review assignment (e.g., that he fails to form the requisite intention) and let $D$ be the set of background facts that are salient to the case less those in $C$ (e.g., that Procrastinate has received an invitation to review a book, that he is the best person for the job, that he has the requisite abilities, that his existing obligations leave him with enough time to complete the job, that he has certain general character flaws). Now let $V$ be the set of worlds that are consistent with the facts in the union of $C$ and $D$ and let $W$ be the set of worlds that are consistent with the facts in $D$. Notice that $V$ is a proper subset of $W$ — the set of worlds at which Procrastinate fails to complete the assignment is a proper subset of the set of worlds that are consistent with the background circumstances specified in $D$.

Now we restate the standard pair of judgments: Procrastinate ought to refuse the assignment given $V$ because he refuses at the highest-ranked worlds at which he fails to complete; he ought to accept (and complete) the assignment given $W$ because he accepts and completes at the highest-ranked worlds in $W$. Put this way, we can see how natural it is to think that the former directive "evaporates", giving the latter directive some sort of primacy. The (concessive) directive to refuse is specified relative to $V$ and, thus, only obtains at the worlds in $V$; whereas the (nonconcessive) directive to accept and complete is specified relative to $W$ and, thus, obtains at the worlds in $W$, *including the worlds in $V$*. Thus, the noted asymmetry: the nonconcessive obligation obtains at the worlds in $V$ (in virtue of those worlds also being members of $W$), while the concessive obligation fails to obtain at the worlds in $W$ that are not in $V$; when we move across worlds in $W$ from those in $V$ to those not in $V$, the concessive obligation ceases to obtain.

To further illustrate the proposed model, consider a case where directives do *not* appear hierarchically arranged: the original trolley problem alongside the variation involving pushing someone from a bridge. Assume (for illustrative purposes) that one is required to pull the lever to save the five in the original case but one is required to refrain

from pushing the person from the bridge to save the five in the variation. It's not at all plausible to declare one directive (e.g., "pull the lever") as subordinate to the other (e.g., "don't push the person"). On the proposed nesting model, this is readily explained by the fact that there is no set of possible worlds that encompasses both scenarios, relative to which one of the directives is specified. There is, of course, a set of worlds that encompasses both cases; namely, the set of worlds at which: five people are tied to a track; a runaway trolley threatens them with imminent death; and you are in a position to prevent the deaths. But neither "pull the lever" nor "don't push" is specified relative to this encompassing set of worlds. Any directives for a particular trolley problem are specified relative to proper subsets of this encompassing set. One subset (call it $R$) includes all and only those worlds at which you can save the five by diverting the trolley to the one; the other (call it $S$) includes all and only those worlds at which you can save the five by pushing one onto the track. Since neither $R$ nor $S$ is a subset of the other, neither "pull the lever" nor "don't push" obtains at all the worlds in the union of $R$ and $S$.[9]

By way of explanation, then: I conjecture that directive principles appear to be arranged hierarchically when the sets of possibilities relative to which they are specified are nested, in the sense modeled above. (That directives are specified relative to particular sets of possibilities is asserted by the Uncontroversial Thesis.[10]) Accordingly, a set of directive principles is maximally superordinate if it is specified relative to a maximally encompassing superset of possibilities. (Which is to say the following mouthful: a maximal directive is an element of a set of principles that codifies the core institutional and behavioral features of the subset of worlds that best realize certain specified evaluative criteria within the maximally encompassing superset of possibilities.)

## 5.  MAKING SENSE OF MAXIMALLY SUPERORDINATE DIRECTIVES

If the proposed nesting model offers a compelling explanation of hierarchically-arranged directives, then a promising avenue for making sense of the notion of a fundamental

---

9   We might nonetheless think that there is a more general directive that obtains at all worlds in $R \cup S$ (indeed, all worlds in the encompassing set). Perhaps "Don't treat persons as mere means" or "Save the greatest number" are candidates. Then the more specific directives are applications of, and thereby subordinate to, the more general directive. That doesn't undermine (indeed, it affirms) my point here: that directives appear to be arranged hierarchically *with respect to each other* when the sets of possibilities with respect to which they are specified are nested.

10   I note, for the sake what is to come, that Estlund also seems to endorse the Uncontroversial Thesis. For instance: "It is true that the non-concessive question contemplates principles *chosen for the way they would operate in an imaginary utopian environment*" (this volume, XX; emphasis added). Additional passages corroborate this.

directive principle is by appeal to the notion of a maximally superordinate directive. We've already seen how to make sense of a maximally superordinate directive for well-defined, small-scale cases, such as Professor Procrastinate. According to the nesting model, "Accept and complete the review" articulates Procrastinate's fundamental moral obligation (granting this for the sake of argument) because it is specified relative to a set of possible worlds that encompasses those at which he fails to complete the review. For these kinds of cases, we might say that the maximally superordinate directive is the *locally maximal directive*.

The question to be considered in what remains is whether we can make sense of the notion of a *globally maximal directive*—that is, a maximally superordinate directive principle not just for a particular well-defined case, but for a suitably wide range of situations of interest. For instance, a globally maximal justice directive identifies not simply which institutions are to be realized (from the standpoint of justice) given some particular factual context; but which institutions are to be realized *tout court*. Put differently, a globally maximal justice directive must make sense of the claim that some institutional scheme is "fundamentally" required by justice, not simply that certain institutional schemes should be implemented in the face of specific (more or less restrictive) feasibility considerations.

Estlund's "global prime requirement" (see chapter in this volume) provides a promising basis for making sense of globally maximal directives.[11] The global prime requirement is an answer to a maximally nonconcessive question: "What principles for the basic structure of society would work out well if there were (contrary to fact) full compliance with them, and *also (contrary to fact) full compliance with all moral standards—full moral compliance?*" (this volume, XX, original emphasis) This can be readily restated in the terms of the optimization model: Which principles characterize the core institutional and behavioral features of the optimal worlds according to certain evaluative criteria, given that everyone fully complies with all (public and private) moral standards?[12] This question contrasts with a set of more concessive questions, which inquire about the principles that characterize the core institutional and behavioral features of the morally optimal worlds given some level of noncompliance with some subset of moral standards.

There are two reasons to be optimistic about using this notion of a global prime

---

11   G.A. Cohen's notion of a "fundamental normative principle" might also be thought appealing in this regard. Space restrictions prevent me from discussing this option in detail. The short reply is that Cohen's fundamental principles are most plausibly thought to perform an evaluative rather than a directive function (see, e.g., Cohen, 2008, 251–253, 265–266, and esp. 348–349). Thus, they cannot help us make sense of fundamental *directive* principles. I consider the idea of fundamental evaluative principles below.

12   Notice that Estlund apparently accepts the optimization model (though not under that guise); see his remarks on constructivist theories of justice (this volume, XX).

requirement to make sense of globally maximal directives. First, the global prime requirement is meant to serve a directive purpose; it putatively performs the deontic function of specifying constraints on action, not simply evaluative criteria (cf. Estlund, 2014). Second, Estlund claims that the global prime requirement bears a kind of primacy over more concessive directives; the global prime requirement is said to be *maximally superordinate* (this volume, XX). As I've already intimated, Estlund argues for this primacy claim by analogy to the Professor Procrastinate case. Procrastinate faces both a nonconcessive directive ("accept and complete") and a directive that concedes to his failure to complete ("refuse the assignment"). Estlund asserts the asymmetry noted above: the nonconcessive directive obtains at all worlds including the worlds at which he fails to complete the assignment; whereas the concessive directive obtains only at the worlds at which he fails to complete the assignment—it "evaporates" once we move outside the subset of worlds at which he fails to complete. The global prime requirement, being the maximally nonconcessive justice directive, bears an analogous primacy over concessive justice directives: the global prime requirement is alleged to obtain across all relevant possible worlds (e.g., those consistent with the actual laws of nature), including partial compliance worlds; whereas partial compliance directives "evaporate" once we move to the set of possibilities consistent with full moral compliance.

With respect to the Procrastinate case, I've proposed that we understand the hierarchical relationship among directives in terms of nested sets of possibilities. The nonconcessive directive gains primacy in this (well-defined, small-scale) case because it is specified relative to a (locally) maximally encompassing set of possibilities. For the global prime requirement to bear an analogous kind of primacy—thereby enabling us to make sense of the notion of a fundamental directive principle—we must make sense of a globally maximally encompassing set of possibilities. This is the task ahead. (To foreshadow: I think we will be led to a relatively unappealing view of fundamental directives.)

We start by noting that, on one plausible way of posing the nonconcessive question to which the global prime requirement is the answer, the global prime requirement is specified relative to a set of possibilities that excludes all worlds involving partial compliance. This is because full compliance is *inconsistent* with partial compliance. To wit, suppose we pose the nonconcessive question thus: Which principles characterize the core institutional and behavioral features of the subset of optimal worlds, given a superset of possible worlds at which everyone fully complies with all (public and private) moral standards? Posed in this way, the question presupposes a set of possibilities that rules out partial compliance. Thus, the global prime requirement is specified relative

to a set of possibilities that does not encompass partial compliance possibilities. Given the nesting model above, we cannot say that the global prime requirement is maximally superordinate because it is not specified relative to a maximally encompassing set of possibilities. In particular, we cannot say that the global prime requirement obtains at worlds involving partial compliance.

This initial worry is readily overcome. Let $W$ be the set of metaphysically and nomologically possible worlds. Let $L_i$ $(i = 0, 1, 2, \ldots, N)$ be a possible level of moral compliance.[13] We arrange these levels in ascending order, with $L_0$ denoting complete non-compliance and $L_N$ denoting full moral compliance. To ensure nested sets of possibilities, we let $W_i$ be the set of worlds at which some level of compliance less than or equal to $L_i$ obtains. For example, $W_5$ is the set of worlds at which some level of compliance less than or equal to $L_5$ obtains; since the level of compliance denoted by $L_4$ is less than $L_5$, $W_5$ encompasses $W_4$.

Since full compliance represents maximum compliance, $W_N = W$, so the set including full compliance possibilities is maximally encompassing.[14] Since the global prime requirement is specified relative to $W_N$, it follows that the global prime requirement obtains at all possible worlds. Thus far, we can say that the global prime requirement is the globally maximal directive according to the proposed nesting model of directive principles.

The key obstacle arises once we appreciate that full moral compliance is consistent with myriad mechanisms for securing full moral compliance. To grasp the point, let's define a new variable to measure an individual's "personal virtue". An individual is *virtuous* to the extent that she conscientiously acts in accordance with the moral directives to which she is subject (even at the expense of any conflicting interests she might have) in the absence of external incentives, monitoring, or enforcement. An "angel" is completely virtuous (cf. Kavka, 1995); a "(conditional) scoundrel" is completely void of virtue — in the absence of external incentives, monitoring, or enforcement, he never complies with moral directives unless doing so happens to align with his self-identified interests. Of course, an individual's degree of personal virtue can lie somewhere between these two

13 Presumably, there are infinitely many possible levels of partial compliance, plausibly uncountably many. I assume that, if uncountable, the set of levels of partial compliance can be partitioned into countably many (and, thus, discretely indexed) equivalence classes, such that levels of partial compliance within the same equivalence class can be treated equivalently for analytical purposes (i.e., the differences among members of an equivalence class do not affect the specification of directive principles).

14 For $W_N$ to be maximally encompassing, it's sufficient to let the union of all $W_i$ for $i = 0, 1, \ldots, N-1$ be a proper subset of $W$ while still identifying $W_N$ with $W$. Then all $W_i$ are encompassed by $W_N$ without the other $W_i$ being nested all the way down, so to speak.

extremes.

It is evident that the mechanisms required to realize full moral compliance depend on the distribution of virtue within a population. At one extreme, full moral compliance is realized by a population composed entirely of scoundrels only if an effective system of external incentives, coupled with reliable monitoring and enforcement mechanisms, is implemented. At the other extreme, full moral compliance is voluntarily forthcoming in a population of angels without the use of external incentive schemes or enforcement mechanisms.[15] There are an infinite number of population-level distributions of virtue between these two extremes, requiring different institutional schemes (crudely, external incentive schemes, coupled with monitoring and enforcement mechanisms) to realize full moral compliance.[16]

Each of these institutional schemes affects the realization of basic moral and social values like liberty, equality, security, community, and so on. The institutional schemes required to realize full moral compliance for distinct distributions of virtue engender distinct realizations of these basic values. To illustrate the point starkly: An institutional scheme required to realize full compliance among scoundrels will tradeoff (e.g.) liberty, equality, welfare, and security differently than an institutional scheme that is sufficient to realize full compliance among angels. (I think it's safe to say, at the very least, that the former institutional scheme cannot afford to grant individuals the same liberties or the same level of personal privacy to realize a similar level of security against, e.g., fraud.) As a result, these institutional schemes manifest distinct normative priorities.

Since the "complete scoundrel" and "complete angel" distributions of personal virtue define distinct (indeed, disjoint) sets of institutional possibilities for realizing full moral compliance, it follows that the set of directive principles specified relative to the institutional possibilities for a population of scoundrels differs (perhaps quite widely) from the set of directive principles specified relative to the institutional possibilities for angels. Yet both sets of directive principles are specified *given full moral compliance.* Thus, we seem to have at least two nonconcessive questions yielding two global prime requirements: one for scoundrels and one for angels. In fact, given the many different distributions of virtue between these two extremes, we have many distinct prime requirements: one for each population-level distribution of virtue.

One might reply that a plurality of global prime requirements causes no trouble for

15   None of this implies that a society of angels has no need for governing institutions. As Kavka (1995) points out, a state is required to resolve reasonable disagreements among angels and to coordinate their attempts to act collectively.

16   As above, we can assume that, if uncountable, the set of population-level distributions can be partitioned into countably many equivalence classes for analytical purposes.

our attempt to make sense of fundamental directive principles. A plurality of fundamental directives is consistent with the existence of fundamental directives (see Estlund, this volume, XX). But this reply faces trouble. The trouble is not that we are left with an infinite set of global prime requirements (although that's likely true). The trouble is that any set of directives specified for a population other than a society of angels *concedes to certain nonideal facts about human behavior.* For instance, the prime requirement relative to the total scoundrel distribution concedes to the distribution of motivational deficiencies found in a population replete with (conditional) scoundrels. The prime requirement relative to this distribution is specified by characterizing the core institutional and behavioral features of the full compliance scoundrel worlds.[17] Thus, the scoundrel prime requirement characterizes a balance of evaluative criteria that is constrained by concessive facts pertaining to scoundrels' motivational deficiencies. Similarly for prime requirements specified relative to less scoundrelly but still not completely virtuous populations; all such prime requirements concede to nonideal facts concerning the distribution of motivational deficiencies within the population. We are thus left with a family of prime requirements, all but one of which are fundamentally concessive.

The only global prime requirement that is fundamentally nonconcessive is the one specified relative to a completely virtuous population — a society of angels. If fundamental directives are identified with maximally nonconcessive directives, then our present attempt to make sense of the notion of a fundamental directive principle apparently leads us to a set of directives suited to a society of angels.[18]

## 6.   ASSESSING THE RESULTS

Suppose for a moment that one finds this conclusion implausible and wishes to deny that the fundamental demands of justice for us — flawed human beings at the actual world — are given by a standard specified for a society of angels. Given the nesting model,

---

17   As an aside, I note that the full compliance scoundrel worlds need not be the normatively optimal scoundrel worlds according to certain specified evaluative criteria. The mechanisms needed to secure full compliance among scoundrels will likely require noteworthy sacrifices of (e.g.) liberty. We can conceive of a reasonable set of evaluative criteria that ranks scoundrel worlds that sacrifice full compliance for the sake of greater liberty above the full compliance worlds. I set aside further discussion of this point, though it can be pressed to create further trouble for an attempt to make sense of "fundamental" directives by appeal to something like Estlund's notion of a global prime requirement.

18   Note that Estlund explicitly denies that the global prime requirement is a "standard for angels". But his remarks at this point suggest that an "angel" is someone who not only complies with all moral directives even in the absence of incentives, monitoring, and enforcement, but also (regularly?) performs supererogatory action (see this volume, XX). I'm not sure how much this difference is supposed to matter.

one can deny this by denying that the angel prime requirement is specified relative to a maximally encompassing set of possibilities; in particular, that it does not encompass the actual world. Yet this ultimately defeats our present attempt to make sense of fundamental directive principles.

To illustrate the point precisely, we adopt a model similar to the one above. Let $W$ be the set of metaphysically and nomologically possible worlds. Let $D_i$ ($i = 0, 1, 2, \ldots, N$) be a possible population-level distribution of virtue. We arrange these distributions in ascending order, with $D_0$ denoting the scoundrel extreme and $D_N$ denoting the angel extreme. (Hence, e.g., the distribution denoted by $D_5$ is closer to the scoundrel extreme than that denoted by $D_6$.) Let $D_\alpha$ denote the distribution of virtue at the actual world, with $\alpha$ equal to some $i < N$. Finally, let $W_i$ be the set of worlds in $W$ such that there are no worlds at which $D_j$ obtains for all $j > i$; that is, $D_i$ is the "highest" distribution of virtue to be realized in $W_i$.

Now suppose that the set of possibilities relative to which the angel prime requirement is specified is not a maximally encompassing set of possibilities. Then $W_N$ excludes at least some of the worlds in $W$ at which non-angel distributions of virtue obtain. At one extreme, $W_N$ includes only those worlds at which $D_N$ obtains; at the other, it includes all worlds except those at which $D_0$ obtain. (We choose $D_0$ without loss of generality.) In any case, $W_N$ is not identical to $W$; so $W_N$ is not maximally encompassing. Given the nesting model, the angel prime requirement specifies the fundamental directives only for those worlds included in $W_N$. If $W_0$ is not encompassed by $W_N$, then the angel prime requirement does not obtain as a fundamental directive for the complete scoundrel worlds. The prime requirement for the complete scoundrel worlds is thus a set of directive principles that concedes to certain facts about humans' moral deficiencies. We can now see how to deny that the angel prime requirement specifies the fundamental demands of justice to which we are subject at the actual world: we deny that $W_N$ encompasses $W_\alpha$, the set of worlds at which $D_\alpha$ — the distribution of virtue at the actual world — obtains.

Now we're in a position to clearly see the trouble this presents for our attempt to make sense of fundamental directive principles. To put a sharp point on it, suppose no $W_i$ is encompassed by another $W_j$ (with $i \neq j$); then each distribution of virtue comes with its own set of fundamental directives. But then it turns out that "fundamental" directives are fundamentally constraint-relative, per the optimization model: "fundamental" directives are solely fundamental relative to a particular set of institutional possibilities given a particular population-level distribution of virtue. All we have in the end is a collection of constraint-relative directives. We are left without anything like a maximally superordinate set of directives. This leaves the notion of fundamental directive principles without any of

the sense of hierarchy or primacy expressed by the Nonconcessive Refrain with which we started this chapter.[19]

In sum, if we suppose that the angel prime requirement is not specified relative to a maximally encompassing set of possibilities, then we are forced to deny the claim that there are maximally superordinate directive principles. But, given the nesting model, this is the claim we're required to vindicate if we are to make sense of the notion of a fundamental directive principle at issue here. Thus, if the angel prime requirement is not specified relative to a maximally encompassing set of possibilities, then we are forced to deny that there are any fundamental directive principles that can sustain the Nonconcessive Refrain.

So the angel prime requirement is our only option for making any sense of the relevant notion of a fundamental directive principle. Let's be precise about this: Our present attempt to make sense of the relevant notion of a fundamental directive principle leads us to the view that the fundamental demands of justice are identified by a set of principles that characterize the core features of an institutional scheme that is sufficient to realize full moral compliance in a society of angels — that is, a society of individuals who conscientiously adhere to the moral directives to which they are subject even in the absence of external incentives, monitoring, or enforcement.

Some might accept this conclusion without pause. But I suspect many will find this a surprising — and relatively unappealing — result. There are at least two intertwined reasons we might find this result unappealing: that a society of angels seems evidently infeasible, and that a standard for angels seems evidently irrelevant. Regarding the infeasibility of realizing a society of angels, we can't demonstrate this of course (see Wiens, 2015*b*); but it seems Panglossian to suppose otherwise. Realizing a society of angels requires not only that everyone fully complies with the moral directives to which they are subject, but that they do so even when those directives conflict with one's other (non-moral) interests and in the absence of external incentives, monitoring, or the threat of sanctions for non-compliance. To be clear, judging a society of angels infeasible does not rule out the possibility that humans can fully comply with the moral directives to

---

19   As a technical aside: We might forge a middle ground, where, for all $i, j$ with $i < j$, $W_i$ is encompassed by some $W_j$ (except when $i = N$) while $W_j$ does not encompass all $W_i$. This means that there are no maximally encompassing sets of possibilities and that the fundamental directives for each $W_i$ are specified relative to some more encompassing $W_j$. But I have trouble seeing how this avoids the problem. First, when $j < N$, we need to tell some principled story for why $W_i$ is encompassed by $W_j$ and not $W_{j+1}$. The most plausible candidate is to say that, given $D_i$, realizing $D_j$ is feasible but realizing $D_{j+1}$ is not. But, second, whatever story we tell here, we are left with the same upshot: per the optimization model, the "fundamental" directives for worlds in $W_i$ are specified relative to the set of salient (e.g., feasible) institutional possibilities. But then there are no globally maximal directives, just a collection of constraint-relative directives.

which they are subject (although I suspect some will remain skeptical about the feasibility of that less demanding possibility too). It solely rules out the possibility of realizing full moral compliance without the use of a system of external incentives, coupled with monitoring and enforcement mechanisms. If one grants that directive principles are subject to some sort of feasibility requirement,[20] then a set of directive principles suited to a society of angels is excluded for imposing a standard suited for an infeasible society.

Tying in the second point, I suspect many theorists of justice think that even an ideally just human society requires a scheme for incentivizing, monitoring, and enforcing compliant behavior. This is evidenced by the widespread agreement among political philosophers that limitations on individuals' motivation to cooperate with others ("limited benevolence") are among the background circumstances that frame judicious theorizing about the demands of justice. At the very least, then, any ideally just society implements an institutional scheme for incentivizing, monitoring, and enforcing compliance for the sake of supporting willing individuals' efforts to comply in the face of limited motivation to do so.[21] As the angel prime requirement expresses normative priorities that are insensitive to the need for institutional mechanisms to support full compliance, it must be irrelevant for any society that requires such institutional supports.

Despite these skeptical remarks, I grant that some might nonetheless endorse the conclusion that our fundamental obligations are specified by principles for a society of angels. (I leave it to proponents to further motivate this view.) For those who find this view implausible, I close by gesturing at a straightforward way to reconcile the idea of "fundamental normative principles" with the denial of fundamental directive principles. The key lies in our distinction between evaluative and directive principles. The foregoing argument suggests that we face serious difficulties trying to make sense of fundamental directive principles so as to vindicate the Nonconcessive Refrain that "justice ought not concede to the facts". But it remains open to us to treat fundamental normative principles as *evaluative* principles: we simply say that fundamental normative principles are those principles that specify basic criteria for comparatively evaluating and ranking any options with respect to each other. That is, fundamental normative principles specify evaluative criteria that are suitable not only for ranking options included in a specific subset of possibilities, but for ranking options across some maximally encompassing set of possibilities (e.g., all metaphysically and nomologically possible worlds).[22] The

---

20  But see Estlund (2011); Gheaus (2013); for my replies, see Wiens (forthcoming) and Wiens (2014) respectively.

21  Rawls's remarks on the role of state institutions in supporting the stable realization of a just society are apt here; see especially Rawls (1999, 237–238, 505).

22  Provided that the set of all options is totally comparable. As the issue of incommensurability or

key point, though, is that fundamental normative principles perform a fundamentally *evaluative* function.[23]

If we take this route, then there are no fundamental directive principles — such principles are fundamentally constraint-relative. Our obligations are thus always specified by reference to a particular set of feasible alternatives. Yet there can be fundamental evaluative principles to guide the specification of our constraint-relative directive principles. I conjecture that many will find this view relatively more appealing than the one according to which our fundamental obligations are specified by directives for a society of angels.

## REFERENCES

Anderson, Elizabeth S. 1999. "What is the Point of Equality?" *Ethics* 109(2):287–337.

Carens, Joseph. 1981. *Equality, Moral Incentives, and the Market.* Chicago: University of Chicago Press.

Charlow, Nate. forthcoming. Decision Theory: Yes! Truth Conditions: No! In *Deontic Modals*, ed. Nate Charlow and Matthew Chrisman. Oxford: Oxford University Press.

Cohen, G. A. 1989. "On the Currency of Egalitarian Justice." *Ethics* 99(4):906–944.

Cohen, G. A. 2008. *Rescuing Justice and Equality.* Cambridge, MA: Harvard University Press.

Cohen, G. A. 2009. *Why Not Socialism?* Princeton: Princeton University Press.

Cowen, Tyler. 2007. "The Importance of Defining the Feasible Set." *Economics and Philosophy* 23(1):1–14.

Estlund, David. 2011. "Human Nature and the Limits (If Any) of Political Philosophy." *Philosophy & Public Affairs* 39(3):207–237.

Estlund, David. 2014. "Utopophobia." *Philosophy & Public Affairs* 42(2):113–134.

Estlund, David. forthcoming. Prime Justice. In *Political Utopias*, ed. Kevin Vallier and Michael Weber. New York: Oxford University Press.

non-comparability creates complications without raising trouble, I set it aside here.

23 If I'm right that Cohen's "fundamental normative principles" perform an evaluative rather than directive function (see footnote 11), then it turns out that Cohen can help us after all. This need not commit us to the further claim that fundamental evaluative principles are fact-insensitive, as Cohen argues.

Gheaus, Anca. 2013. "The Feasibility Constraint on the Concept of Justice." *The Philosophical Quarterly* 63(252):445–464.

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

Goldman, Holly S. 1976. "Dated Rightness and Moral Imperfection." *Philosophical Review* 85(4):449–487.

Graham, Peter A. 2011. "'Ought' and Ability." *Philosophical Review* 120(3):337–382.

Jackson, Frank and Robert Pargetter. 1986. "Oughts, Options, and Actualism." *Philosophical Review* 95:233–255.

Jacopo Uberti, Luca. 2014. "Good and Bad Idealizations in Political Theory." *Theoria* 80(3):205–231.

Kavka, Gregory. 1995. "Why Even Morally Perfect People Would Need Government." *Social Philosophy and Policy* 12(1):1–18.

Kolodny, Niko and John MacFarlane. 2010. "Ifs and Oughts." *Journal of Philosophy* 107(3):115–143.

Kratzer, Angelika. 1991. Modality. In *Semantics: An International Handbook of Contemporary Research*, ed. Arnim von Stechow and Dieter Wunderlich. Berlin: de Gruyter pp. 639–650.

Lawford-Smith, Holly. 2013. "Understanding Political Feasibility." *The Journal of Political Philosophy* 21(3):243–259.

Portmore, Douglas. 2009. "Consequentializing." *Philosophy Compass* 4(2):329–347.

Rawls, John. 1999. *A Theory of Justice*. 2 ed. Cambridge, MA: Harvard University Press.

Sen, Amartya. 2002. *Rationality and Freedom*. Cambridge, MA: Harvard University Press.

Valentini, Laura. 2009. "On the Apparent Paradox of Ideal Theory." *Journal of Political Philosophy* 17(3):332–355.

Vranas, Peter B. M. 2007. "I Ought, Therefore I Can." *Philosophical Studies* 136:167–216.

Wiens, David. 2014. "'Going Evaluative' to Save Justice From Feasibility—A Pyrrhic Victory." *The Philosophical Quarterly* 64(255):301–307.

Wiens, David. 2015*a*. "Against Ideal Guidance." *Journal of Politics* 77(2):433–446.

Wiens, David. 2015*b*. "Political Ideals and the Feasibility Frontier." *Economics and Philosophy* 31(3):447–477.

Wiens, David. forthcoming. "Motivational Limitations on the Demands of Justice." *European Journal of Political Theory* .