

CONFESSIONS OF A UTOPOPHOBE*

BY MICHAEL HUEMER

Abstract: Ideal theorists in political philosophy seek to describe a perfect political society, and to evaluate political principles by reference to their consequences in a world where everyone complies with the principles. I argue that ideal theory is not needed to set goals for practical inquiries, nor to define justice, nor to enable rankings of injustices. Nor is it useful to theorize about very different kinds of society that might occur in the far future. Ideal theory tempts us to make each of three kinds of error: it tempts us to propose norms that no specific agent can act on, to posit crazy exaggerations of moral virtues, and to place too much trust in abstract philosophical reasoning. A better approach to normative questions is to rely on analogical arguments starting from uncontroversial intuitions about concrete scenarios.

KEY WORDS: ideal theory, utopia, justice, Rawls, intuitionism, particularism

I. A THEORY OF TRANSPORT

Two philosophers, Ida and Nora, were out for a drive when they heard a bang and Ida started to lose control of the car. When they pulled over and got out, Nora pointed at the left front tire.

Nora: That tire looks flat. I'm going to change it.

Ida: Whoa! Slow down there. Changing the tire *might* be the way to go. But we can't know that until we first answer some theoretical questions.

Nora: Really? Like what?

Ida: First we need to determine the nature of the Perfect Car, so we can know what changes to our car will count as improvements.

Nora: It's getting dark. Maybe we should change the tire now, then we can talk about perfect cars while we're on our way home.

Ida: But until we identify the Perfect Car, we have nothing to aim at.

Nora: Well, I don't think the perfect car would have a flat tire.

Ida: Yes, now you're starting to see the importance of ideal theory. But even if it turns out that the flat tire is a problem, we don't know whether it would be a more or less serious problem than the dust on the body. We can't know that until we have a clearer picture of the ideal car.

* I would like to thank David Schmitz, David Estlund, and the other contributors to this volume for their many smart, interesting, and helpful comments on the issues in this essay. I have done my best to address those I knew how to address.

- Nora: Alright, tell me about your ideal car.
 Ida: Well, as we've agreed, the perfect car would never get a flat tire.
 Nora: I didn't say that. I don't see how a car can be made so it *never* gets a flat.
 Ida: You're getting confused. We're not discussing what's *feasible* now, we're discussing what's *desirable*. Now, since the Perfect Car would be incapable of getting a flat tire, it must be some sort of hovercraft, maybe with antigravity technology . . .
 Nora: Since we don't have antigravity technology, I think we should stick with using the wheels.
 Ida: Don't be such a defeatist! Many great technological advancements have been made in history.

No one reasons like Ida in everyday, practical contexts. Nevertheless, some of the motivations for engaging in "ideal theory" in political philosophy are uncomfortably reminiscent of Ida's arguments above.

Of course, the immediate practical problem facing Ida and Nora renders a detour into theoretical reflection of any sort ill timed for them, but that is not the focus of my critique. The problem that I want to point to in the above dialogue concerns the specific content of the arguments Ida gives: those are unsound arguments, and their unsoundness may shed light on the unsoundness of parallel arguments in political philosophy.

In what follows, I argue (i) that political philosophy does not need ideal theory, and (ii) that ideal theory is often harmful to the goals of political philosophy. My view is not that there is no place for ideal theory, but that its utility has been overstated, while less theoretical and less idealized approaches are often underappreciated. But first, a few words about the notion of ideal theory.

II. WHAT IS IDEAL THEORY?

Ideal theory is not a theory, but a *way of theorizing*. Sometimes, ideal theory is taken to be concerned with the ideal society generally — that is, a society that embodies all social values. Call this "general ideal theory." More commonly, ideal theory is spoken of as concerned specifically with *justice*, which is but one value among many. Call this "ideal justice theory."

What is ideal justice theory? In contemporary literature, two things are commonly said. One is that ideal (justice) theory aims to describe a perfectly just society, regardless of whether such a society is attainable. The other is that ideal theory aims to identify the principles of justice that would be best to adopt if all individuals were to comply fully with these principles. Both accounts appear in Rawls:

The reason for beginning with ideal theory is that it provides, I believe, the only basis for the systematic

grasp of these more pressing problems. [. . .] At least, I shall assume that a deeper understanding can be gained in no other way, and that the nature and aims of a perfectly just society is the fundamental part of the theory of justice.¹

[I]n assessing conceptions of justice the persons in the original position are to assume that the one they adopt will be strictly complied with. The consequences of their agreement are to be worked out on this basis.²

I shall refer to these two versions of ideal theory, respectively, as “perfection theory” and “strict compliance theory.”³

Though they are often spoken of as equivalent, strict compliance theory is importantly different from perfection theory. Strict compliance theory, as Rawls describes it, involves a controversial methodological commitment not demanded simply by the project of describing a perfectly just society. What Rawls proposes in the second quotation above is that one should decide what the principles of justice are, partly on the basis of an assessment of what the consequences would be if certain principles were generally complied with. One could reject this methodology without rejecting the project of theorizing perfection. Suppose, for example, that I am a libertarian intuitionist, and I believe that human beings have certain natural rights that can be known by ethical intuition. I might then think that a perfectly just society would be one in which those rights are never violated, and I might think it interesting to describe such a society. Beliefs about the consequences of perfect respect for rights need not play any role in my selection of the principles describing our natural rights, nor need I have engaged in any Rawlsian or other constructivist project.

What goes for ideal justice theory goes equally for general ideal theory: we can distinguish the *strict compliance* version of general ideal theory from the *perfect society* version. In general strict compliance theory, we select normative social principles in part on the basis of the consequences of everyone’s following those principles. In general perfection theory, we try to describe a society that perfectly embodies all social values.

¹ John Rawls, *A Theory of Justice*, rev. ed. (Cambridge, MA: Harvard University Press, 1999), 8.

² *Ibid.*, 126.

³ Cf. Laura Valentini’s three-way distinction among strict compliance theory, utopian theory, and end-state theory (“Ideal vs. Non-ideal Theory: A Conceptual Map,” *Philosophy Compass* 7 [2012]: 654–64). My “perfection theory” is a combination of Valentini’s “utopian theory” and “end-state theory.”

III. FOUR IDEAL THEORISTS

Before turning to my objections to ideal theory, let's review some examples of this type of theorizing.

A. *Rawls*

Rawls's theory of justice is an example of ideal justice theory, both in the strict compliance sense and in the perfectionist sense. Rawls argues for his two principles of justice on the ground that these principles would be chosen by the parties in his Original Position thought experiment.⁴ In making this choice, the parties, by Rawls's stipulation, assume that the principles will be met with strict compliance once adopted. This assumption is essential to the reasoning. For example, the parties are said to favor the Difference Principle because this minimizes their risk — that is, it makes the worst outcome they might face as good as possible — on the assumption that the government will faithfully implement the principle.

What if we dropped the assumption of strict compliance? The parties might then worry that government leaders would not know how to maximize the welfare of the least advantaged citizens. Or that implementation of the chosen policies would be marred by bureaucratic incompetence. Or that legislators and government agencies would fall under the sway of interest groups who cynically advance their own interests under the guise of aiding the poor. Once we introduce realistic concerns such as these, it is no longer clear that adopting the Difference Principle minimizes risk. It might be better to adopt some principle that is easier for leaders to apply and harder for rent-seekers to manipulate. Rawls's assumption of strict compliance therefore makes a great difference to his theory of justice.

Rawls contends that the perfectly just society would be one in which his principles of justice were perfectly followed; thus, what he offers us is also an ideal justice theory in the perfectionist sense.

B. *Cohen*

In *Why Not Socialism?* G. A. Cohen offers a general perfection theory. He asks us to imagine a friendly camping trip in which a group of campers all voluntarily share resources (pots and pans, coffee, canoes, and so on),

⁴ Here are the two principles: "FIRST PRINCIPLE: Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all. SECOND PRINCIPLE: Social and economic inequalities are to be arranged so that they are both: (a) to the greatest benefit of the least advantaged [. . .] and (b) attached to offices and positions open to all under conditions of fair equality of opportunity" (Rawls, *Theory of Justice*, 266). Part (a) of the second principle is "the Difference Principle."

and all voluntarily contribute to the various tasks that need to be done (fishing, cooking, cleaning, and so on). Everyone treats everyone else as an equal, and no one expects special privileges because of their superior talents, luck, or heredity. This scenario, says Cohen, exhibits the key socialist values of community and equality. The ideal society would be one in which we could somehow extend the mode of interaction of this camping trip to the entire society. He admits that we do not know how to do this, but he thinks this no reason to disparage the ideal, nor should we give up trying to attain it.⁵

C. *Carens*

Joseph Carens develops a theory designed to show how a society could achieve the socialist ideal of perfect equality without sacrificing economic productivity. Carens's proposal is that everyone should be taxed in such a way that everyone's after-tax income is equal; however, all should nevertheless voluntarily strive to produce as much pretax income as possible, out of a sense of social duty. Call this system (including the stipulation about the altruistic behavior of citizens) the "Carens Market." How could this be brought about? Carens thinks we can socialize people such that they derive the same satisfaction from performing their social duty that people presently derive from increasing their personal, disposable income.⁶

D. *Brennan*

Lest it appear that ideal theory is the exclusive province of the political left, Jason Brennan offers a capitalist ideal. Brennan imagines a capitalist society in which public-spirited citizens voluntarily cooperate and trade with each other, always respecting each other's rights, but always willing to help others in need through private charity.

Part of what Brennan is doing in *Why Not Capitalism?* is constructing a parody of Cohen's *Why Not Socialism?* But Brennan is also serious about his ideal theory: he thinks that describing utopia is a worthwhile project,

⁵ G. A. Cohen, *Why Not Socialism?* (Princeton, NJ: Princeton University Press, 2009), 80, 82.

⁶ From Joseph Carens, *Equality, Moral Incentives, and the Market* (Chicago: University of Chicago Press, 1981), 96: "The basic position I shall adopt [...] is that all human motivation is the result of socialization and that it is theoretically possible to socialize people in the egalitarian system into placing as much value on the satisfactions associated with performing their social duty to earn pre-tax income as individuals in the PPM [private-property market] system place on the satisfactions derived from acquiring income for consumption." Carens does not claim that his system would make for a perfect society, since he thinks it retains some undesirable features of capitalism (Carens, *Equality, Moral Incentives and the Market*, xi, 178). Nevertheless, it seems fair to consider this as a kind of ideal theory.

and that the correct version of utopia is capitalist.⁷ The reason is that in the ideal capitalist society, a variety of lifestyles would be available, catering to the great variety of kinds of people who exist.⁸ People who want to live on a kibbutz could do so, while those who want to compete in the business world would be free to do that, and so on.

IV. MISGUIDED MOTIVES FOR IDEAL THEORY

A. "Ideal Theory Sets the Goal"

Why engage in ideal theory? Some argue that ideal theory is a necessary foundational part of normative social theory. One argument is that, as Rawls puts it, "until the ideal is identified . . . nonideal theory lacks an objective, an aim, by reference to which its queries can be answered."⁹ Stemplowska and Swift concur:

[W]ithout knowing our long term goal, a course of action that might appear to advance justice . . . might nonetheless make less likely, or perhaps even impossible, achievement of the long-term goal. . . . [Ideal theory] tells us where we are trying to get to in the long run.¹⁰

As a justification for engaging in perfection theory, this argument is unconvincing; it begs the central question at issue by assuming that normative political theory must pursue perfection. The sensible critic of ideal theory will not claim to pursue the perfect society while having no idea of what that would be like. The sensible critic will say: we need not, and perhaps *should* not, conceive ourselves as pursuing perfection. Perhaps we should aim only to satisfy. Or perhaps we should aim to solve particular, circumscribed problems — say, injustices in current immigration policy, or inequities in the treatment of racial minorities.¹¹

⁷ Jason Brennan, *Why Not Capitalism?* (New York: Routledge, 2014), chap. 4. In fairness, Brennan's utopia is more realistic than the others since it does not require dramatic shifts in most people's motivational structure.

⁸ Compare Robert Nozick's "framework for utopias" (*Anarchy, State, and Utopia* [New York: Basic Books, 1974], chap. 10).

⁹ *The Law of Peoples* (Cambridge, MA: Harvard University Press, 1999), 90. For similar remarks, see Rawls, *Theory of Justice*, 8; Ingrid Robeyns, "Ideal Theory in Theory and Practice," *Social Theory and Practice* 34 (2008): 341–62, at 344–45.

¹⁰ Zofia Stemplowska and Adam Swift, "Ideal and Nonideal Theory," in *The Oxford Handbook of Political Philosophy*, ed. David Estlund (Oxford: Oxford University Press, 2012), 379. For similar remarks, see A. John Simmons, "Ideal and Nonideal Theory," *Philosophy and Public Affairs* 38 (2010): 5–36, at 34.

¹¹ Thus, Charles Mills's well-known criticisms of ideal theory are driven by the desire to see racial inequities addressed, which he thinks is made less likely by the focus on ideal theory ("'Ideal Theory' as Ideology," *Hypatia* 20 (2005): 165–84).

Why might one rest content with such modest aims, rather than pursuing perfection? Perhaps because the modest aims seem more achievable, or because they are simply more salient to the particular theorist at a particular time. Rawls and his allies do not present an argument that the political philosopher must take perfection as his goal; they simply assume this.

Compare the case of Ida and Nora from Section I above. It is false that without an ideal of the Perfect Car, Nora has no goal to aim at. Her goal is to change the tire and get the car moving again. This is a modest aim, but no less genuine for that. Life is full of such modest aims.

I anticipate an objection: as Stemplowska and Swift suggest above, action aimed at addressing a particular, limited social problem might turn out to conflict with the ultimate attainment of the perfect society. We cannot know whether this is the case until we have done some amount of ideal theory. If it turns out to be the case, then we should perhaps desist from that action. Thus, it might be said, we must do some ideal theory before addressing particular social problems.¹²

But compare the analogous reasoning in the flat tire example. Nora proposes to solve the car's salient problem by changing the tire, without first engaging in ideal theory. Ida protests that until they determine the nature of the Perfect Car, they cannot know that changing the tire won't *somehow* prevent them from attaining vehicular perfection; thus, they must engage in ideal theory before changing the tire.

What should be Nora's response? Probably something like this: "Prior to engaging in ideal theory, my initial credence that changing the tire would somehow prevent some much better outcome for the car from coming about is very low — too low, in fact, for it to be worth spending significant resources investigating the possibility. Since you have said nothing about what better outcome might be prevented, nor how it might be prevented, your comment leaves my credence unchanged. Thus, I still plan to change the tire."

Similarly, suppose that I am reasoning about immigration policy. I have an argument that nearly all restrictions on immigration are unjust, with restrictions on migration from poor countries to wealthy countries being especially harmful and unjust.¹³ This argument rests on ethical intuitions about cases; it does not rest on a theory about the perfectly just society. Should I withhold judgment on the issue, or hold off from advocating relaxed immigration laws, on the grounds that easing immigration restrictions *might somehow* prevent perfect justice from being attained someday? With no concrete reason to think that this would be the case, and no account of how it would be the case, the answer is no.

¹² See also Simmons, "Ideal and Nonideal Theory," 22–24.

¹³ See my "Is There a Right to Immigrate?" *Social Theory and Practice* 36 (2010): 429–61.

This is not to deny that easing immigration restrictions may create other problems, or that the theorist should address those problems. The point is simply that objectors must identify specific alleged problems; the general possibility of the policy's interfering with ultimate justice in some unknown way, based on our ignorance of the nature of perfect justice, is not a potent objection.

B. *"Ideal Theory Defines Justice"*

The defenders of perfection theory might find my arguments question-begging, for I have assumed that we *can* pursue incremental improvements in justice or other political values without having an ideal of perfection — and this is what they deny. Injustice, they might say, is simply that which diverges from the perfectly just society; without knowing what the perfectly just society is like, we have no way of judging whether any particular action or state of affairs constitutes an injustice. As Simmons says, "[t]o dive into nonideal theory without an ideal theory in hand is simply to dive blind, to allow irrational free rein to the mere conviction of injustice and to eagerness for change of any sort."¹⁴ There are two ways of reading this: First, perhaps Simmons is suggesting that without a *theory* of justice, any putative identification of a particular injustice must be an irrational mere conviction. Second, perhaps he is suggesting that any rational assessment of the justice of a particular action or state of affairs must be based upon a belief about perfect justice, perhaps the belief that the particular action or state either is or is not consistent with perfect justice.

Again, compare the case of Ida and Nora: Nora has no vision of the perfect car, nor does she have a comprehensive theory about the norms governing cars or movement in general. This does not mean that her belief that the flat tire is a problem is an irrational mere conviction, nor is she diving blind into changing the tire. It is not true in general that we make judgments about concrete cases in the light of general theories, nor do we typically evaluate things by reference to a standard of perfection. Consider two more examples:

- i) Someone asks me what I think of the television series *Game of Thrones*. I have no general account of aesthetic merit or even entertainment value. I have read not one article in aesthetics, nor have I any idea what the leading theories in the field are. I do not know what a perfect television series would be, if such a thing even makes sense, and in this, I suspect, I am in the same boat with the experts in aesthetics. Yet despite my appalling ignorance, I say: "It's a good show."

¹⁴ Simmons, "Ideal and Nonideal Theory," 34. Cf. Brennan, *Why Not Capitalism?* 71 (approvingly explaining Cohen's defense of utopian theorizing): "[I]f you imagine a society in which people sometimes did wrong things, you'd be imagining a society with some injustice in it, and thus be imagining a less than fully just society. So, if you care at all about what justice requires, you have to ask what utopia would be like."

- ii) Someone asks me who I think was a better human being: Mohandas Gandhi, or convicted serial murderer Ted Bundy? I have no general, abstract theory of virtue. I do not know what a perfect human being would be, if such a thing is even possible. Nevertheless, I hazard the judgment that Gandhi was better than Bundy.

These examples are not anomalies; they are perfectly normal cases of human judgment. They illustrate that evaluations — whether absolute or comparative, aesthetic or moral — are typically made without reference to any standard of perfection. Neither of these examples is of a mere guess or a blind conviction. I know that *Game of Thrones* is a good show, and I know that Bundy was worse than Gandhi. I would in fact consider these judgments to place *constraints* on acceptable general theories: any acceptable theory of aesthetic merit or of virtue must accommodate judgments like these.

If, therefore, the proponents of ideal theory wish to maintain that evaluations of justice depend upon a theory of perfection, they owe us a substantial argument for this conclusion. Why should assessments of justice differ so dramatically from assessments of aesthetic merit, virtue, and so on?

C. “Ideal Theory Enables Comparisons”

Another argument on behalf of ideal (perfection) theory is that one needs a conception of the perfect society in order to make comparative judgments — judgments assessing, for instance, which of two injustices is more serious. In Rawls’s view, one judges this in part by the prioritization of principles provided by ideal theory. For example, because the ideal theory of justice prioritizes personal liberty over equitable wealth distribution, violations of liberty rights are more serious than instances of inequitable distribution.¹⁵ More generally, one might think that the seriousness of an injustice is a matter of how far the unjust practice or outcome deviates from the practices or outcomes of the perfectly just society.

My response to this argument should be clear from the remarks above: the argument rests on an epistemological error. When I judge Bundy worse than Gandhi, I do not do this by reference to a theory of the perfect human; I have no such theory. When I judge *Game of Thrones* a better television drama than *Baywatch*, I do not do so by reference to an account of the perfect TV drama. When Nora judges the flat tire to be a more urgent problem than the dust on the car’s body, she does not do this by reference to a theory of the perfect car. There is no reason to think that comparative evaluations in general are made by reference to a vision of perfection, and no reason has been given for thinking that comparative justice judgments would be special in this respect.

¹⁵ See Simmons, “Ideal and Nonideal Theory,” 34: “[T]he priority to be given to grievous (over less grievous) injustices is to be understood in terms of the lexical ordering of the principles of ideal theory violated by the injustices at issue.”

D. "Today's Utopia Is Tomorrow's Reality"

Another reason for engaging in ideal theory is that what today seems like a fantasy may one day become quite realistic. It therefore makes sense to aim high. As Brennan observes, throughout most of human history, deaths due to violence were vastly more common than they are today. If someone a thousand years ago described a society like ours, they would likely have been dismissed as utopian. Perhaps similar dramatic advances, that today appear overly utopian, will be made in the future.¹⁶

I share Brennan's optimism: we will one day live much better than we do now. Nevertheless, I think utopian theorizing is largely useless.

Compare Ida's hypothesis, in the flat tire story, that we may one day develop antigravity technology. This might be true, but this possibility has no practical bearing on Ida and Nora's trip. Antigravity technology is not available to them presently, and they do not know how to make it available. You might think that their present behavior might nevertheless affect their prospects for attaining antigravity technology at some distant future time. But Ida and Nora have no idea whether antigravity will ever be available, nor do they know even in broad outline how it would work. In that context, the speculation that something they do today might somehow help or hinder future antigravity efforts is idle. Discussion of such a possibility does not help them solve any problem or pursue any goal that they have any rational way of pursuing.

The same is true of G. A. Cohen's speculations about a future society in which everyone interacts like people on a friendly camping trip. We do not have such an option available to us now, we have no idea whether it will ever be available, nor do we know even in broad outline how it would come about. In that context, discussion of this possibility is idle. It does not help us solve any problem or rationally pursue any goal.

I think the future holds great promise: our distant descendants will live much better than we, due not only to technological and economic advancements, but probably to institutional and cultural advancements as well. But we have no way of predicting these developments and no rational way of trying to either help or hinder them. Imagine the position of someone in 1500 A.D., trying to plan for the world of 2016. There is no reasonable way that they could have guessed even roughly what our society would be like. As a result, if anyone in 1500 A.D. made plans for the twenty-first century, those plans are surely worthless. Because the pace of change has dramatically advanced in the past few centuries, the future is even less predictable to us than our world was to the denizens of 1500 A.D. Thus, any plans we might make for the next age are almost surely worthless. The point here is that Brennan's general observation about the

¹⁶ Brennan, *Why Not Capitalism?* 71–72.

likelihood of progress — even if true — does not give us good reason for engaging in utopian theorizing.

There is no sharp dividing line between possibilities that are too remote to be usefully entertained and those that are close enough to be worthwhile. My claim that Cohen's utopia is idle speculation is thus a judgment call: in my judgment, a society in which socialism would work would have to be so far from our society that we cannot now productively discuss that society or plan for its arrival.¹⁷

E. *The fate of strict compliance theory*

The above arguments focus on the theory of the perfect society or perfect justice. In slogan form, the three arguments we have considered claim that (i) ideal theory sets our goal, (ii) ideal theory defines justice, and (iii) today's utopia may be tomorrow's reality. All of these are claims about the necessity or value of theorizing about the perfect society. Notably absent is any defense of strict compliance theory, in the sense defined in Section II. That is, none of these is an argument that the correct normative political principles should be identified by asking what results each candidate principle would have in a world of perfect compliance with all officially recognized political principles. That methodological claim seems to have sneaked in the door with the similar-sounding but distinct claim that we should aim at achieving a perfect society.

There is room for debate about the role of consequences in assessing political principles. But *if* we are to take account of consequences, why would the relevant consequences be those that would result from the adoption of a principle in a world of strict compliance with all officially adopted principles? Wouldn't the relevant consequences be those that would result from the adoption of a given principle in a world like ours?

Some ethicists endorse *rule consequentialism*, the theory that the right thing to do is always to comply with the best system of moral rules, where the best system of rules is the one that would have the best consequences if it were universally practiced.¹⁸ Something like this idea is occasionally invoked in ordinary moral reasoning when one asks, "What if everyone did that?" For example, you should not walk across a newly planted lawn, even though your doing so will have no discernible impact on the lawn, because if everyone in a similar position behaved in a similar manner, the lawn would be ruined. (But note that one need not endorse a pure rule

¹⁷ One might wonder why my own defense of anarchism (*The Problem of Political Authority* [New York: Palgrave Macmillan, 2013], part 2) is not similarly too utopian. The answer is that the anarcho-capitalist system does not require alterations of human nature; it works with normal levels of human selfishness and strife — or so I argue.

¹⁸ See Richard Brandt, *Morality, Utilitarianism, and Rights* (Cambridge: Cambridge University Press, 1992), chap. 7.

consequentialist ethic to find this reasoning compelling — one might simply hold that, if it would be bad if everyone did x , this counts as a reason against doing x .) Perhaps strict compliance theory is motivated by similar thoughts.

But whatever one thinks of walking across lawns, there seem to be more directly relevant analogies that show the error of strict compliance theory. Strict compliance theory in political philosophy concerns the choice of political or legal principles that are to be officially recognized and enforced by the state. And we do not typically think that such principles should turn on the consequences of strict compliance, as opposed to the realistically expectable consequences of a policy. Consider two examples:

(i) Should there be a rule according to which judges must recuse themselves from presiding over cases involving their own family members — for instance, so that one cannot serve as the judge in a trial where one's own wife is the defendant? Suppose we evaluate this rule by its effects in a world of perfect compliance with all recognized principles of justice. In such a world, judges will always be perfectly fair and objective, even when their own family members are on trial. Thus, one might argue, there is no need for a recusal rule.

Now, there may be other reasons why the recusal rule is a good idea, but leave that aside. The point here is that the preceding reasoning for the claim that we do not need a recusal rule is completely unconvincing, precisely because of the way it relies on the assumption of perfect compliance. Much more compelling is the argument that since in fact judges cannot be expected to be objective when their own family members are on trial, no judge should sit in such a case. This is a good argument; it is surely not to be evaded by our simply stipulating perfect compliance with all principles of justice.

(ii) What should be the nation's drug policy? Assume that the harms caused by recreational drug use outweigh both the enjoyment drug use brings and the freedom that is lost when these drugs are outlawed. If this is so, it might seem to create a justification for having the state outlaw these drugs.

That justification, however, seems to turn on the assumption that drug laws will be by and large obeyed. In fact, such laws face massive noncompliance problems. In the United States, the consequences of this noncompliance include (a) the annual expenditure of billions of dollars on enforcement; (b) increased rates of theft and violent crime; (c) increased police corruption; (d) pressure for the government to erode civil liberties; and (e) the incarceration of close to half a million people, severely harming both the prisoners and their families.¹⁹

¹⁹ See my "America's Unjust Drug War," in *The Right Thing to Do: Basic Readings in Moral Philosophy*, 5th edition, ed. James Rachels and Stuart Rachels (New York: McGraw-Hill, 2010), 223–36; and William Chambliss, "Another Lost War: The Costs and Consequences of Drug Prohibition," *Social Justice: A Journal of Crime, Conflict and World Order* 22 (1995): 101–24.

Now imagine a theorist responding to these problems caused by prohibition by insisting that what matters is what consequences the policy would have in a world where everyone complied with it. Surely this would be wrong; what matters is the consequences that can realistically be anticipated.

So, in both the judge recusal case and the drug policy case, the correct political or legal norms are not to be identified on the basis of predicted consequences in a world of strict compliance. If ideal theorists believe that principles for the basic structure of society are somehow different, such that they *should* be chosen based on their consequences in a world of strict compliance, these theorists owe us a justification for that belief.

V. UTOPIAN ILLUSIONS

A. *Agentless norms*

Thus far, I have only criticized arguments in favor of ideal theory. I turn now to ways in which ideal theory is often philosophically misguided.

The first problem is that perfection theorists are often led into enunciating *agentless norms*, that is, statements to the effect that something ought to be done, where there is no agent to whom the ostensible imperative might plausibly be directed. I think all such claims are either false or nonsensical: predicates such as “ought” and “should” relate an agent to a possible action; it cannot be that *x* ought to be done, unless there is someone who ought to do it.

Typically, these agentless norms are nominally attached to *society* (note: not merely the state); that is, it is said that society ought to be a certain way, to adopt some principle, or to bring about some result.²⁰ David Estlund speaks of what “all people, together” should do.²¹ Joseph Carens’s theory stipulates not only the policies to be adopted by the state, but also how the public is to react to those policies — it is that entire package, including the policies plus how people react, that Carens is advocating.

To whom might Carens’s recommendation be addressed? It cannot be addressed to the state, since the state lacks the power to implement it. The state could implement *its part* of Carens’s system (the redistributive taxation scheme), but doing so would be disastrous since the other part of the system (that citizens maximize pretax income) would not occur. The state therefore *should not* do its part in the Carens Market. David Estlund is right to note that the fact that we are unwilling to do *x* does not in general

²⁰ See David Estlund, “Human Nature and the Limits (If Any) of Political Philosophy,” *Philosophy and Public Affairs* 39 (2011): 207–37, at 235–37.

²¹ *Democratic Authority: A Philosophical Framework* (Princeton, NJ: Princeton University Press, 2008), 266. Estlund (“Human Nature,” 217) also recognizes Carens’s system as a valid normative political theory.

cast doubt on the claim that we ought to do *x*, however intransigent our refusal may be.²² But the unwillingness of *other* agents to comply with some plan *does* cast doubt on the idea that we should pursue it. In this case, the likely reaction of the public explains why the state should not attempt to implement the Carens Market, nor should the state do its part in the Carens Market, nor should the state do anything else in the vicinity.

Nor can Carens's theory plausibly be addressed to any individual citizen: it is not the case, for example, that *I* should bring about the Carens Market, since I cannot do so; nor should I *try* (futilely) to bring the system about; nor should I *do my part* in the system, etc. The reason for this cannot lie in predictions of my own behavior (in this I am in agreement, again, with Estlund) — it does, however, lie partly in rational predictions about the behavior of others. I should not attempt to implement the Carens Market, since I know that I would not succeed in radically changing the behavior of the rest of the agents in my society.

Carens's theory only makes sense, then, if it can be addressed to everyone simultaneously, the state conjoined with all the individual citizens. But any putative norm addressed to society, or to all people together, or to the state conjoined with the citizens, is an agentless norm: neither society, nor all people, nor the state conjoined with its citizens, constitutes an agent.

Note that I am not imposing a strict individualist constraint, to the effect that only *individuals* can have obligations, reasons for action, and so on. I am not denying the existence of collective agency, nor am I denying that it can make sense to speak of what a group ought, collectively, to do. The state, for example, can be viewed as an agent; the state performs actions, has obligations, and so on. Similarly, the Exxon Mobil Corporation, the Catholic Church, and the American Philosophical Association are all agents.

But not just any collection of people constitutes an agent. What exactly is required to have a collective agent, I do not know, but it seems that the individuals must at least have in place some reasonably robust mechanisms for coordinating with each other. These might include formal decision-making processes for an organization to which they belong, or even informal norms in some cases. Thus, a very small society — say, a primitive tribe, or a kibbutz — might be able to attain enough cohesiveness to constitute an agent.

But surely a collection of millions of people, almost all of whom are complete strangers to each other and have no effective mechanism of coordination, does not constitute an agent. It therefore does not make sense to speak of what such a collection should do. For instance, it makes no sense to speak of what all Americans, together, ought to do.

Now, you might think that the members of a society *do* have the requisite coordination mechanism, provided that they have a democratic government.

²² Estlund, "Human Nature," 207–14.

I find this proposal dubious, since I do not think that the state (even if democratic) can legitimately speak for society. Be that as it may, even if we regard the existence of the state as establishing collective agency for *society*, we must still recognize limits on this agency determined by the limitations of the state. In other words, if the way that society collectively acts is through the state, then what society can bring about is limited to what the state can bring about. As we have seen above, Carens's proposal cannot be brought about by the state, nor can any proposal that demands strict compliance.

I have no objection to a normative theory addressed to the government, no matter how unlikely it may be that the government will take heed. I likewise have no objection to a theory addressed to individuals — even to *all* individuals (that is, to each and every individual) — no matter how unlikely most individuals are to take heed. My objection is that Carens's theory is not addressed to the government, nor to any individual, nor to any other agent; for that reason, it cannot be a correct normative theory.

B. Crazy standards

There is a second way in which ideal theory tends toward false norms. The emphasis on describing perfection tempts theorists toward postulating extreme versions of norms that in their moderate versions are widely accepted. For instance, generosity and concern for the good of society are virtues, while selfishness is a vice. Ideal theorists are therefore tempted to posit an extreme version of altruism as the moral ideal, one in which an agent has no more concern for himself than he has for a person completely unknown to him. I think that is a mistake. It is not just that I think we cannot make people be that way; I do not think that such a person would be an ideal human being. Rather, such a person would be crazy.

This point is dramatized in an episode of the television series *House, M.D.*²³ A team of doctors is trying to diagnose patient Benjamin Byrd, who has been brought into the hospital after an unexplained episode of fainting. Byrd also happens to be a successful businessman who has given away almost all of his money to charity. While in the hospital, Byrd meets another patient who needs a kidney transplant, whereupon Byrd offers the other patient one of his own kidneys. Dr. House proposes that whatever Byrd's condition is, it must have a neurological component. Dr. Adams demurs: there's nothing wrong with the patient's brain, she says; he is just very generous.

²³ David Shore (writer), Sara Hess (writer), and Greg Yaitanes (director), "Charity Case" (television series episode), *House, M.D.*, season 8, episode 3, aired Oct. 17, 2011 (NBCUniversal Television Distribution).

Doctor Hadley (also known as “Thirteen”) then goes to speak with Byrd. She mentions (falsely) that she has polycystic kidney disease and she needs a transplant. Byrd immediately offers her his kidney, whereupon the following dialogue ensues:

- Hadley: But you already promised it to someone else.
 Byrd: I have another one. Saving one life is good; saving two is better.
 Hadley: You give away two kidneys and you die.
 Byrd: I can live on dialysis for years.
 Hadley: Yes . . . and then you die.
 Byrd: And then I can donate my other organs. Heart, lungs . . . I could save four or five more lives.

Hadley retreats slowly from the room, then calls another doctor and reports, “This guy is crrrrazy.”²⁴

Why would such a patient be crazy, rather than merely very virtuous? The reason cannot merely be that his preferences are statistically very unusual. Rather, I think the reason is something like this. Human beings have a range of instincts and emotions, which form the core of the human motivational system. The normal range is very wide but not unlimited. The degree of self-sacrifice displayed by Byrd in that scene goes beyond what is consistent with a well-functioning human motivational system. Healthy human beings sometimes sacrifice their lives for the good of others, but such sacrifices are driven by love for particular others, not by abstract appreciation for the total utility of the universe. The only plausible ways for an individual with otherwise good life prospects to wind up motivated to give away his heart, lungs, and kidneys to strangers are (a) if that individual has a psychiatric condition (as in the *House* case), or (b) if the individual has a radically different motivational system from us, divorced from the human instincts and emotions with which we are familiar. In either case, I do not think such a being is one we could sensibly hold up as an ideal for ourselves.

Extreme levels of either altruism or selfishness are pathological in humans. A less extreme (but still high) degree of altruism is admirable, and of selfishness, execrable. In between, there is a wide range of merely normal motivation. *Pace* Peter Unger, you are not a terrible person if you fail to donate most of your income to charity, even if that really would be the objectively best thing to do.²⁵ The danger of ideal theory is that it tempts us to set up a pathological ideal of motivation and then to unfairly condemn ourselves, both as individuals and as a species, for failing to satisfy

²⁴ *Ibid.*, at 32:44–34:26.

²⁵ I refer to Peter Unger’s *Living High and Letting Die: Our Illusion of Innocence* (Oxford: Oxford University Press, 1996).

this unreasonable ideal. This self-condemnation does not make us better; it only makes us dispirited.

C. *The dangers of abstraction*

Ideal theory in political philosophy commonly relies on very abstract philosophical reasoning. Rawls is a case in point: he bases his central argument in *A Theory of Justice* on highly abstract claims about appropriate constraints on reasoning about principles of justice, plus some abstract normative intuitions, such as the unfairness of resource distribution based on morally arbitrary characteristics.²⁶ Such theories should be viewed with suspicion from the start, because abstract philosophical reasoning in general tends to be wrong. Throughout the history of the subject, when a philosopher thinks that she has a compelling abstract proof of a theory, the theory is almost always false. We know this because on most issues, philosophers have defended many incompatible views. *Ex ante*, therefore, there is no reason to think that developing an abstract theory of justice would cause us to make more reliable judgments about which policies or institutions are just. Simmons worries that acting without a theory is like diving blind. But we should rather worry that acting with a philosophical theory is like diving into a hallucination.

A critic might object that I am relying upon abstract philosophical reasoning, even as I argue against relying upon abstract philosophical reasoning. The argument of the preceding paragraph putatively supports the general philosophical theory that general philosophical theories tend to be mistaken.

In reply, note that I am not advancing that claim as a starting premise; I have come to the conclusion that abstract, general philosophical theories tend to be unreliable on the basis of concrete empirical evidence that could hardly be mistaken: the history of disagreement about philosophical theories. Furthermore, my conclusion here need not be established with a high degree of confidence, since all I am trying to do is *cast doubt on* conclusions arrived at in ideal theory. If you think there is even a *good chance* that ideal theories are typically false, that gives you reason to doubt conclusions based on such theories.

But what is the alternative? Should we give up trying to discern the demands of justice? Should we give up philosophical research in general?

I do not know the best thing to say about philosophy in general. But for political theory, I suggest that we may develop a more reliable methodology than has hitherto been generally practiced. This superior methodology would be one founded on essentially uncontroversial intuitions about particular cases, rather than on whatever general, abstract principles a

²⁶ Rawls, *Theory of Justice*, 11–19.

particular theorist finds appealing. Of this, I shall have more to say in the following section.

VI. A PLEA FOR NONIDEAL NONTHEORY

A. *Nonideal nontheory*

In Rawls's taxonomy, there are two parts to political philosophy: ideal theory and nonideal theory. Ideal theory tries to describe perfect justice or the perfect society. Nonideal theory, in Rawls's conception, aims to figure out how we can approach the outcome described by ideal theory.

I want to enter a plea for a third approach to questions of justice and other social values; call it "nonideal nontheory." In nonideal nontheory, rather than trying to describe a perfect society, we try to say what policies and institutions should actually be adopted. We need not claim that these policies or institutions will ever bring about utopia; we merely claim that they are probably superior to the available alternatives (this is the "nonideal" part). And rather than appealing to a general, abstract account of the principles of justice or other social values, we appeal to intuitions about particular cases (this is the "nontheory" part).

For example: earlier I alluded to my stance on immigration restrictions. Here, briefly, is why I think these restrictions are unjust. Suppose that I meet a person named Marvin who is trying to reach a marketplace to buy some food, which he desperately needs. And suppose that I intentionally, forcibly prevent Marvin from reaching that marketplace, with the result that he starves. Absent special circumstances, this action would be very wrong on my part. I am not inferring this from a general principle of wrongness, rights, or justice; it just seems obvious on its face. But immigration restrictions seem analogous: they forcibly prevent needy people from reaching a place where they could satisfy their needs by trading with willing partners. Unless there is some crucial disanalogy that I have overlooked, then, immigration restrictions are seriously wrong.

Of course, there is more to say about this argument, and one must consider candidate disanalogies between the Marvin example and immigration restrictions. But that is enough to give a sense of how I think one may identify unjust policies and institutions; similar arguments can be given with respect to other issues.²⁷ Notice that the argument proceeds without anywhere mentioning the nature of perfect justice; notice, too, that the key normative premise is an essentially uncontroversial intuition about a concrete case.

²⁷ On immigration, see my "Is There a Right to Immigrate?" *op. cit.* I apply the approach to other issues in my "America's Unjust Drug War," *op. cit.*; "Is There a Right to Own a Gun?" *Social Theory and Practice* 29 (2003): 297–324; and *The Problem of Political Authority*, *op. cit.*

A few guidelines about appropriate concrete cases: what is important is that the audience have definite intuitions about the case, that the intuitions of different individuals generally agree, and especially that the intuitions not be tied to the reader's particular political ideology. Typically, one must cite a metaphysically possible case, but one need not cite a case that is likely to occur or that has ever in fact occurred, provided that the audience has definite intuitions about it. The familiarity of a case may of course affect our ability to muster definite intuitions about it, as will the specificity of its description. In accordance with the general motivation for relying on intuitions about cases in the first place, it is important that cases be described relatively concretely. Thus, in the example discussed above, Marvin is walking to a local marketplace to buy food. The example would lose effectiveness if it were merely given as one in which "agent *A* seeks access to resource *R*."

This approach to political theorizing has obvious affinities to moral particularism, the view that there are no exceptionless moral principles and that moral judgments must be made on a case-by-case basis.²⁸ Particularists should find my approach to political philosophy attractive. Nevertheless, I am not hereby *committed* to moral particularism, because I do not claim that exceptionless moral principles *do not exist*. I claim only that we do not now know a sufficient number of informative, general moral principles to reliably assess the justice of policies and institutions on that basis, nor are we in a position easily to acquire such knowledge; at the present, therefore, the more reliable approach is to base our assessments on analogies to particular cases about which we have firm intuitions.

B. *Are intuitions reliable?*

Some hold that ethical intuitions about cases are unreliable because these intuitions often appear to be inconsistent or influenced by morally irrelevant factors.²⁹ In addition, since intuitions, by definition, are not based on arguments, there is reason to worry that they may simply reflect a particular theorist's prejudices. On the subject of justice especially, intuition may simply reflect one's political ideology.

In reply, let me first explain why the central alternative approach provides no solution to these worries; I will then briefly explain how the concerns may best be addressed.

If we seek to avoid reliance on ethical intuitions about cases, the alternative methodology for arriving at normative conclusions would be to rely

²⁸ See Jonathan Dancy, *Moral Reasons* (Oxford: Blackwell, 1993).

²⁹ See Peter Singer, "Ethics and Intuitions," *Journal of Ethics* 9 (2005): 331–52; Unger, *op. cit.*; Joshua Greene, "From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology?" *Nature Reviews Neuroscience* 4 (2003): 847–50. For general doubts about ethical intuition, see also Walter Sinnott-Armstrong, *Moral Skepticisms* (Oxford: Oxford University Press, 2006).

upon more abstract, theoretical claims. But if examples of misguided intuitions about *cases* support skepticism about such intuitions, then wouldn't examples of misguided *abstract theories* support skepticism about abstract theories? Mistaken concrete intuitions tend to occur in certain kinds of difficult cases, such as those involving conflicts between consequentialist and deontological reasons for action, not cases like the starving Marvin example discussed above. Meanwhile, by all accounts, the overwhelming majority of abstract philosophical theories that have ever been seriously advanced are false, including those whose proponents considered the theories to be conclusively proven. Surely the track record for abstract philosophical theorizing looks worse than the track record for ethical intuitions about cases.

As to the concern about intuitions being influenced by ideology, appeal to abstract theorizing again offers no remedy.³⁰ People who start from different ideological orientations vary drastically in their assessments of the merits of abstract philosophical arguments — it is not as though, for example, libertarians who are exposed to Rawls's theory of justice are typically convinced; nor are left-wing theorists who are exposed to Nozick's rights-based arguments. Abstract normative theories, including most of the products of ideal theory in political philosophy, fail to overcome biases; they merely serve as vehicles for the expression of one's biases.

How, then, should we address concerns about the reliability of particular case intuitions? I have discussed this issue elsewhere;³¹ here I will just briefly state that the best approach is to identify the most common kinds of biases leading us astray in our judgments about cases, and to avoid use of particular intuitions that show signs of such bias, while continuing to rely on those intuitions for which we lack any specific grounds for doubt. This approach shows sensitivity to the challenges but avoids throwing the baby out with the bathwater.

To address ideological bias in particular, we should strive, in our reasoning about social issues, to start from ideologically neutral intuitions — that is, intuitions about cases that would be shared by the great majority of people, regardless of their political orientation. For instance, liberals, conservatives, libertarians, anarchists, socialists, and moderates can all agree that I should not forcibly stop Marvin from reaching the marketplace to buy food; this intuition is thus an appropriate starting point for a political argument. It is not open to charges of ideological or other bias.

In addition, we should take care to avoid the pitfalls discussed above (Section V). In developing normative theories, we should limit ourselves to recommendations that can be addressed to specific, existing agents who

³⁰ There are exceptions to this: certain formal ethical intuitions, such as the transitivity of "better than," are immune to charges of ideological bias, as discussed in my "Revisionary Intuitionism," *Social Philosophy and Policy* 25 (2008): 368–92, at 383–87.

³¹ "Revisionary Intuitionism," *op. cit.*

are capable of bringing about what is proposed. We should evaluate policies and institutions based on their expected consequences in a world with realistic levels of compliance. We must take account of the consequences of noncompliance, both on the part of ordinary citizens and on the part of government agents. And we should embrace only those norms that might be followed by reasonably normal, psychologically healthy human beings.

Philosophy, University of Colorado, Boulder