https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

THERE IS NO SUCH THING AS IDEAL THEORY*

By Jacob T. Levy

Abstract: *In this essay, I argue against the bright-line distinction between ideal and nonideal normative political theory, a distinction used to distinguish "stages" of theorizing such that ideal political principles can be deduced and examined before compromises with the flawed political world are made. The distinction took on its familiar form in Rawls and has enjoyed a resurgence of interest in the past few years. I argue that the idea of a categorical distinction — the kind that could allow for a sequencing of stages of theorizing — is misconceived, because wholly "ideal" normative political theory is a conceptual mistake, the equivalent of taking the simplifying models of introductory physics ("frictionless movement in a vacuum") and trying to develop an ideal theory of aerodynamics. Political organization and justice are about moral friction in the first instance. I examine both logical and epistemological arguments for the position that we need the uniquely idealizing assumptions of ideal theory in order to arrive at, or to know, a genuine theory of justice or political morality; and I find them wanting. Such assumptions as full compliance, consensus, and the publicity principle of universal knowledge about consensus can sometimes be useful, if used carefully and with justification; but they are not categorically different from other idealizing and abstracting assumptions in generating normative theory. What is referred to as "nonideal" theory is all that there is, and it is many kinds of theory, not one — the many ways in which we learn about justice and injustice, and seek to answer questions of practical reason about what ought to be done in our political world.*

KEY WORDS: ideal theory, nonideal theory, realist political theory, full compliance, consensus, historical injustice, moral epistemology, circumstances of justice

I. Ideals, Idealizations, and Friction

The now-conventional bright-line distinction between ideal and nonideal normative political theory distinguishes "stages" of theorizing such that ideal political principles can be deduced and examined before compromises with the flawed political world are made, indeed that theorists *ought* to do the first before turning to the second so as to avoid contaminating our aspirations

* This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada. Thanks to Diane Shnier, Elisa Muyl, and Léa Gruyelle for research assistance on this paper, and to McGill's Arts Research Internship Awards for support for Shnier's and Muyl's work. For discussion and comments on earlier drafts I am grateful to audiences at the Association for Political Theory Annual Meeting; the Toronto-area Conference for the Study of Political Thought; the Stanford Political Theory Workshop; the McGill Research Group on Constitutional Studies; a virtual workshop at Brown University's Political Theory Project; and to the other contributors to this volume. I am particularly grateful to Arash Abizadeh, Naomi Choi, Joseph Carens, David Estlund, Megan Rose Blomfield, Jakob Reckhenrich, John Tomasi, Christopher Freiman, Michael Huemer, and Jennan Ismael for comments during and after those discussions, as well as to David Schmidtz for sustained conversation on these topics over several years. Estlund in particular commented on several drafts as the article neared completion.

doi:10.1017/S026505251600025X
© 2016 Social Philosophy & Policy Foundation. Printed in the USA.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

and principles with premature contact with those flaws. The distinction took on its familiar form in Rawls and has enjoyed a resurgence of interest in the past few years. Much of this interest has come from committed defenders of ideal theory.[1] But much has come as well from those who seek a structured and rigorous approach to nonideal theorizing; these theorists find such an approach by accepting some ostensibly ideal theory as taken off the shelf, and by then applying that idea to nonideal circumstances.[2] The spectre of ideal theory haunts even some of the best work about justice in a nonideal world.

In this essay I aim to banish the spectre. In the realm of political philosophy, or of theorizing about justice, there is no such thing as ideal theory.

The idea of a categorical distinction, the kind that could allow for a sequencing of stages of theorizing, is misconceived. The hope for a normative political theory that is ideal in some absolute sense is a conceptual mistake, the equivalent of taking the simplifying models of introductory physics ("frictionless movement in a vacuum") and trying to develop an ideal theory of aerodynamics.[3] Like aerodynamics, political life is *about*

[1] See, in various ways, G. A. Cohen, *Rescuing Justice and Equality* (Cambridge: Harvard University Press, 2008); David Estlund, "Utopophobia: Concession and Aspiration in Democratic Theory," in Estlund, *Democratic Authority: A Philosophical Framework* (Princeton, NJ: Princeton University Press: 2008); Estlund, "What Good Is It? Unrealistic Political Theory and the Value of Intellectual Work," *Analyse und Kritik* (2011): 395–416; Estlund, "Human Nature and the limits (if any) of Political Philosophy," *Philosophy and Public Affairs* 39, no. 3 (2011); A. John Simmons, "Ideal and Non-Ideal Theory," *Philosophy and Public Affairs* 38, no. 1 (2010): 5–36. I take these to be the leading contemporary expositors and defenders of ideal theory.

[2] See, for example, Burke Hendrix, "Where Should We Expect Change in Non-Ideal Theory?" *Political Theory* 41, no. 1 (2013): 116–43; Liam B. Murphy, *Moral Demands in Nonideal Theory* (New York: Oxford University Press, 2000). Some important articles that do not straightforwardly advocate or only advocate one or the other category of theory include Laura Valentini, "Ideal vs. Nonideal Theory: A Conceptual Map," *Philosophy Compass* 7 (2012): 654–64; Pablo Gilabert, "Comparative Assessments of Justice, Political Feasibility, and Ideal Theory," *Ethical Theory and Moral Practice* 15 (2012): 39–56; Pablo Gilabert and Holly Lawford-Smith, "Political Feasibility: A Conceptual Exploration," *Political Studies* 60 (2012): 809–25; Zofia Stemplowska, "What's Ideal About Ideal Theory?" *Social Theory and Practice* 34 (2008): 331–40; Zofia Stemplowska and Adam Swift, "Ideal and Nonideal Theory," in David Estlund, ed., *The Oxford Handbook of Political Philosophy* (Oxford: Oxford University Press, 2012); Charles Larmore, "What Is Political Philosophy?" *Journal of Moral Philosophy* 10 (2013): 276–306. For pieces in the same spirit as the present argument see David Miller, "Justice For Earthlings," in *Justice for Earthlings* (Cambridge: Cambridge University Press, 2012); David Schmidtz, "After Solipsism," unpublished manuscript [currently in press and due to be published in *Oxford Studies in Normative Ethics* 6 (forthcoming, 2017)]; Andrew Sabl, *Hume's Politics* (Princeton, NJ: Princeton University Press, 2012); and, of course, Judith Shklar, *The Faces of Injustice* (New Haven, CT: Yale University Press, 1990).

[3] As best as I can now recall, I arrived at the use of this analogy before seeing David Schmidtz's far more evocative comparison with "ignoring wind resistance in predicting the behavior of a parachute," but the idea is certainly the same (David Schmidtz, "Nonideal Theory: What It Is and What It Needs to Be," *Ethics* 121, no. 4 [2011]: 772–96). For an objection, see Jennan Ismael, "A Philosopher of Science Looks At Idealization in Political Theory," in the present volume. Ismael emphasizes the value of idealization as a mode of inquiry including in physics, but this only shows that one or another idealization is useful for one or another purpose. As will become clear below, I do not deny this; the claim of ideal theory is something much stronger. She briefly concedes that her argument does not provide any particular support for the sequencing of ideal and nonideal that is found in Rawls, and that sequencing is an important part of what I reject here.

friction: no friction, no politics or justice. Or, to take an analogy closer to our disciplinary home: ideal normative political theory is not like microeconomic models with their assumptions of perfect competition and perfect information — radically simplifying assumptions that can be useful in important ways. It is rather like microeconomics with added assumptions of superabundance and the impossibility of scarcity: a confused muddle, because microeconomic theory precisely is a way of thinking about choice among limited goods.

To theorize is to engage in simplification and abstraction — that is, in idealization. And to have a normative vision is to be in at least some small degree idealistic: we imagine that in at least some way, the world might be other than, and better than, it is. But I will argue that the particular idealizations relied on by ideal theory are at best some idealizations one might choose, not those that give us distinctive access to valid normative principles or goals. They are in some respects particularly strange idealizations, especially close to idealizing away friction. That does not mean that they might never be useful, but it does tell against making them crucial in political thought without a powerful reason to do so. I will examine, and reject, the reasons that have been offered by ideal theorists for thinking that these particular idealizations are necessary, indeed so crucial as to justify a sequencing between a prior, categorically distinct, ideal theory that relies upon them and a later, applied, nonideal theory that does not. If those arguments for the *necessity* of ideal theory fail, then so does the justification for sequencing and for categorical distinction, and, thus, so does ideal theory itself.

There are meaningful differences between more and less idealizing theories of justice, and productive debates to be had about which is more useful, and when. But such a spectrum does not require categorical distinctions and a sequence of stages. It is always incumbent on normative theorists to identify which features of the world they take to be reformable, which idealizations they are making, which they do not and are not, and why. The idea of "ideal theory" creates one spot on the spectrum of idealization that is thought *not* to be subject to those questions, and this is illusory. The ideal theories that appear relevant to political life necessarily smuggle in nonideal premises in order to justify the need for politics and justice altogether. Those that fail to do so also fail to be relevant, at best collapsing into a moral theory that lies across an unbridged gap from an articulation of political ideals of justice.

My target in this essay is thus in one sense narrow. It is *categorically* ideal theory: the idea that there is some particular kind of idealization that allows us to sharply distinguish between ideal and nonideal theory, with ideal theory offering privileged access to moral truth in a way that gives it priority over nonideal. I do not mean to criticize *idealization* (in the sense of simplification or abstraction) as such. Nor do I criticize *idealism* in normative theory in any blanket way. Indeed, I will argue that there

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

is no general relationship between the arguments for ideal theory as a kind of idealization and normative idealism, though when they have been conflated I will have some arguments that will impugn some kinds of normative idealism. Finally, and perhaps most importantly, I do not criticize what we might call *idealer* theory. With respect to either idealization or idealism, theories come in degrees of greater and lesser, and with that I have no quarrel. It is claims of categorical distinction, claims of priority, and (typically) combinations of the two that I mean to deny. Through the remainder of the essay, my use of the phrase "ideal theory" will always be restricted to the kinds of theories that claim distinct, prior, and/or privileged status for some particular idealization, and will exclude merely *idealer* theory. (This will also spare me and the reader from any further use of that ugly neologism.)

In another sense, however, the target of my argument is broader than some readers might first assume; it is not limited to Rawlsian theorizing about justice, though that is the most familiar example. Due to their prevalence in the literature, I will direct special criticism at the specific sense of ideal theory used in Rawlsian debates and to the particular idealization of full compliance. I will also discuss the assumption in Rawlsian ideal theory of consensus: in order to think clearly about a theory of political justice, part of what we do is imagine a well-ordered society in which all affirm the same principles of justice. Consensus is implied by full compliance, in a way that is sometimes overlooked; we could hardly all commit to abiding by the same principles of justice if we did not already agree to those principles. But full compliance and full consensus abstract from importantly different features of social life, and warrant separate discussion.

One more idealization from the Rawlsian literature, the absence of historical injustice, will also be relevant to the discussion. But I do not mean to limit the indictment to the Rawlsian type of ideal theory, or indeed to theories of justice. G. A. Cohen's variant of ideal theory is not only a radicalization of Rawls's; it is in some relevant methodological ways a different kind of enterprise. But I do think it commits some of the mistakes I will be discussing herein. And more generally, ideal normative theories of *politics* do as well. The Rawlsian assumption of consensus generalizes to other models that try to derive fundamental normative truths about politics by denying the basic, dissensual, character of political life. The Habermasian school of democratic theory and its offshoots in deliberative democratic theory — indeed any account that derives principles of democracy from an initial idealization of universal agreement — normally offers another instance of ideal theory: democratic politics without the politics, without friction.

This sometimes underacknowledged connection between the Rawlsian demand for consensus under ideal theory and models of deliberative democracy that foreground consensus means that my argument connects the debate about "ideal and nonideal theory" to that about so-called

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

"political realism and moralism."[4] Ideal theory is a subset of political moralism as that is characterized by realists such as Bernard Williams, and by their predecessors such as Hannah Arendt.[5] But much nonideal theory is in that sense moralism, too, in part because it so often accepts the job description of applying an already-settled (and moralistic) ideal theory to the conditions of the nonideal world. In saying that there is no such thing as ideal theory, I am also necessarily denying that there ought to be that kind of nonideal theory. Rejecting the ideal theory assumption of consensus may make it appropriately harder to imagine that the task of nonideal theory is to moralize at unpleasant facts, and may thereby bring the practice of nonideal theory closer to at least some kinds of realism.

II. The Logic of Ideal Theory

The dominant sense of ideal theory at stake in current debates was first developed by Rawls. Rawls understood himself to be *doing* ideal theory, developing a conception of justice that came prior to such problems in the world as a theory of punishment, compensatory justice, or a theory of just war. But he did not describe this only as the exercise that he happened to be engaged in, one choice among many.[6]

> I examine the principles of justice that would regulate a well-ordered society. Though justice may be, as Hume remarks, the jealous, cautious virtue, we can still ask what a perfectly just society would be like. Thus I consider primarily what I call strict compliance as opposed to partial compliance theory . . . The latter studies the principles that govern how we are to deal with injustice . . . 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 insight 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.[7]

[4] On realism, see William Galston, "Realism in Political Theory," *European Journal of Political Theory* 9 no. 4 (2010): 385–411; Bernard Williams, "Realism and Moralism in Political Theory," in *In the Beginning Was the Deed: Realism and Moralism in Political Argument* (Princeton, NJ: Princeton University Press, 2005); Williams, "Political Philosophy and the Analytic Tradition," in *Philosophy as a Humanistic Discipline* (Princeton, NJ: Princeton University Press, 2008). For an argument denying that ideal/ nonideal and moralism/ realism are closely related distinctions, see Enzo Rossi and Matt Sleat, "Realism in Normative Political Theory," *Philosophy Compass* 9, no. 10 (2014): 689–701.

[5] Hannah Arendt, *The Human Condition* (New York: Harcourt Brace, 1958).

[6] Sometimes David Estlund has described his view of utopian theory in that way: he is describing one thing that one might do, defending the legitimacy of one intellectual project that one might have among others. See Estlund, "What Good Is It?" But that is decidedly not Rawls's account.

[7] John Rawls, *A Theory of Justice*, rev. ed. (Cambridge: Belknap Press, 1999), 8, emphasis added.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

It is this belief, these assumptions, that I particularly deny, not the idea that strict compliance might ever be a useful idealization at all.

In considering candidate theories of justice, Rawls supposed, his hypothetical choosers would construe each of them on the assumptions of strict compliance and universal endorsement (as well as on an assumption that they will apply to a society that is not "burdened," including by a *history* of injustice; we will return to that below). They would aim to understand and to choose what justice demands, what the just society would be. Nonideal considerations are eligible to be considered only in a subservient way. In the face of departures from justice *as that was specified at the ideal stage* — in the face of crimes, laxness of character, unwillingness to live up to justice's high standards, unjust laws, uncompensated historic wrongs, or unjust present restrictions on liberty — one should worry about how to best approximate just conditions. The lack of universal compliance with the demands of justice might pose difficulties of implementation that must be addressed by nonideal theory.

But such considerations must *not* be allowed as a consideration in the selection of a conception of justice in the first place. "That's too demanding and too hard" is ineligible as an objection at the stage of identifying acceptable theories.[8] This concern with strict compliance — later increasingly described as *full* compliance — has remained central to the analysis of ideal theory. David Estlund and G. A. Cohen in different ways have emphasized that "I/we *won't*" cannot serve as a reason to reject an understanding of justice, even if it serves as a declaration of intent to put insuperable barriers to justice's implementation. The criminal's bad will — literally, his replacement of *will* by *will not* — might tell us how difficult justice is to attain but must not become an element of the definition of justice itself, lest we lose the ability to call injustice by its right name.[9]

Justice is, as Hume saw and Rawls alluded to in the quotation above, a remedial virtue. Under conditions of superabundance or boundless generosity, "the cautious, jealous virtue of justice would never once have been dreamed of . . . Justice, in that case, being totally useless, would be an idle ceremonial, and could never possibly have place in the catalogue of virtues . . . [I]t seems evident, that the use of justice would, in this case,

[8] In Rawls's own theory, there must be a qualification here for the discussion of the strains of commitment; see discussion in Simmons. If we place a great deal of weight on the strains of commitment, then we might think Rawls himself falls afoul of Estlund's denial that motivational limits are relevant for the evaluation of claims of justice or obligation — that is, that Rawls is not an ideal theorist by the standards of the contemporary literature. I am unsure what to think about this, as a matter of Rawls interpretation.

[9] For an objection, see David Wiens, "Motivational Limitations on the Demands of Justice," *European Journal of Political Theory,* forthcoming, online first DOI 10.1177/1474885115578446. See also David Estlund, "Reply to Wiens," *European Journal of Political Theory,* forthcoming, online first DOI 10.1177/1474885115602369

be suspended by such an extensive benevolence, nor would the divisions and barriers of property and obligation have ever been thought of."[10]

Now, some contemporary theorists of justice deny Hume's account, most conspicuously Cohen, for whom justice is something like a Platonic category of morality, not the mere Humean "rules of regulation" that govern life in the cave.[11] But Rawls does *not* reject Hume on this point (as Cohen complains). This presents a puzzle: How can Rawls be committed to an ideal (in his sense) theory of justice while still relying on Hume's understanding of justice and its circumstances?

A. John Simmons has provided the best-developed answer to this question, saying that "full compliance" is the one *unrealistic* assumption Rawls allows himself in his quest for a "realistic utopia" — the only one that violates the Rousseauian requirement to understand "*men as they are* and institutions as they might be." A theory of justice should be analyzed, so the argument goes, and the choice among such theories made, in the first instance without reference to whether persons will do as they oughtn't and violate the principles. We choose among ideal principles in their best and most coherent forms, without worrying about whether the noncompliant will jeopardize the stability of the system. I think this claim underlies much of the appeal of a priority for ideal theory. It corresponds with some of our normal practices of moral judgment and also (as we will see) the *judicial* judgment from which the ideal of justice analogically extends. I will refer to this as the *logical* defense of ideal theory, to be distinguished from an *epistemological* defense to be discussed later: without the full compliance, publicity principle, and consensus requirements, we logically *cannot* arrive at a conception of justice, as we will have allowed the theory we derive to be corrupted by compromises with bad faith and bad actors.

But, as James Madison said, "If men were angels, no government would be necessary." So-called ideal theories of normative political justice are derived assuming full compliance with whatever they happen to demand, but do nothing to justify being in the business of normative *political* theorizing about *justice* if humans could be assumed to fully comply with whatever morality demands of them. Taking "strict compliance" seriously would mean assuming away the crime that justifies the state's control of the means of violence, the limited beneficence that sits at the base of theories of justice in property and in the coercive provision of social welfare,

[10] Hume, *A Treatise of Human Nature,* III.1. Hume also held that justice pertained only to cases of *moderate* scarcity, and that what we now think of as lifeboat ethics lies outside of justice. I think this is right, but it is not relevant to our concerns for the moment and so I set it aside.

[11] I object to this account in Jacob T. Levy, review of *Rescuing Justice and Equality*, in *Political Theory* 38, no. 4 (2010): 593–96. Justice is the legalistic virtue and the virtue of legal systems; it is *ius*. In my view, Cohen's argument is an attack on the idea of justice, not a rescuing of it. See also Stemplowska and Swift, "Ideal and Nonideal Theory." Compare Allan Beever, *Forgotten Justice: The Forms of Justice in the History of Legal and Political Theory* (Oxford: Oxford University Press, 2013).

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

and more generally the failings that make politics and justice unavoidable. To limit ourselves to the Humean circumstances of justice: there is something deeply paradoxical about deriving rules for the management of our limited beneficence by imagining that *whatever rules we derive* will attract universal cooperation. Why not simply imagine more beneficence? Indeed I think that is what Cohen does: he rejects the circumstances of justice as well as the juridical character of justice so completely that he feels free to articulate an unrelated ideal of fraternal beneficence and simply call it justice for the sake of giving it greater political weight. But Rawls does not; he means the model of choosing a theory of justice from among candidate theories by imagining them all operating with strict compliance and general consensus to be compatible with recognizing the circumstances of justice. Although the method flirts with assuming away the problem of justice by allowing us to imagine that any principles, even universal beneficence, would be met with compliance, Rawls does take for granted that we are choosing among theories of *justice,* filtering out theories like Cohen's at the outset.

Full compliance can be a coherent tool of analysis at the micro level, one transaction or one category of cases at a time. When we are trying to figure out how to split a dinner bill among several people at a restaurant, our calculations should not in the first instance be affected by any thought like "if Jim's portion comes out higher than he likes, he's the kind of person who might slip out the back door and stiff us altogether." In order to determine what each owes, we presume that each will actually pay what we jointly decide they owe. Otherwise the enterprise will be plagued by moral hazard and indeterminacy. Fairness demands analysis in terms of full compliance, even if compromise with actors of bad will enters later.

Similarly, consider the priority of rights over remedies in private law theory — the core, after all, of a theory of Humean justice, as befits the legalistic virtue derived from *ius.*[12] Conceptually, we must know who owns what before we can consider problems of trespass or conversion. The norm of observing contracts precedes a particular breach, and provides the benchmark against which damages for breach are measured. Restitution for a harm aims, as the phrase goes, to make the victim *whole;* it depends on an understanding of the victim's condition before the harm was committed. Rights come first, remedies for their violation second. Universal observance of the rules of justice is taken as a default; offenses against the rules are measured against that background.[13]

[12] Again, see Beever, *Forgotten Justice*. In his terms, I am here endorsing his argument that commutative justice is as fundamental to what justice means as distributive justice, although I have other disagreements with his account.

[13] Even here it may be revealing that most common law developed for a very long time around access to remedies; institutionally, writs came before rights. But of course this is because judicial action was only needed in case of breach or violation. Society was assumed to run on an ongoing basis of most rights being respected most of the time, and so institutional

Within the confines of one domain of rules or another, this seems sound enough. But of course these domains of rules are entirely within the Humean overall framework. In identifying who owns what, we momentarily consider the hypothetical world in which the rules *of ownership* have been universally respected. But the question "why have rules of ownership at all" still gets us to the Humean answer — or, if one prefers, to Kant's "unsocial sociability."

I referred above to what *fairness* demands in the case of the restaurant bill. Rawlsian "justice as fairness" and the analysis of society as an idealized, unified, closed, shared system of social cooperation generalize up from the importance of strict compliance in particular or domain-specific cases — the rules for dividing a bill, the rules of property ownership — to the case of every (social) thing at once.

It is easy to lose sight of this fact, given the Rawls-Cohen debate about the basic structure. It is Cohen rather than Rawls who seems to argue for the isomorphism of macro-social and micro-individual morality, arguing that our intuitions about sharing equally on a camping trip teach us about the basic moral attractiveness of socialism. I think this is misleading. The camping trip notwithstanding, Cohen's theory really has the structure of importing a society-wide rule into the individual soul; it is a scaling *down* from macro to micro, not the other way around. The *ethos* that animates the citizens of a Cohenite just society is one of individual commitment *to the overall social rule*. Rawls, by contrast, genuinely does engage in a kind of scaling up. Fairness demands that the person who cuts the cake not be allowed first choice of a piece; this insight about everyday fairness is foundational to the elaboration of justice *as* fairness.

Whatever the other merits or demerits of justice as fairness, I think that Rawls's theory scales up the assumption of full compliance in an illegitimate fashion, and that this has been the source of much confusion. In modeling the constructivist choice among conceptions of justice, he treats justice as a kind of large-N generalization of fairness, rather than (in Hume's fashion) as a socially distinctive category resting on distinctive macro-level questions. The Humean remedial character of justice, its impersonality, its tight traditional entanglement with juridical and judicial concepts including sanctions, makes it something meaningfully other than personal fairness writ large. Recall that for Hume justice is an "artificial virtue," discontinuous with the natural virtues for precisely these reasons.

---

cognitive energy was better spent on specifying writs than on conducting a comprehensive census of rights. One could therefore plausibly think that the history of the common law is compatible with the thought that (tacitly acknowledged) rights precede (institutionally developed) remedies. And I would not want to place too much weight on the quirky history of the common law. Still, I find the example telling, if nothing else then as evidence against the epistemological defense of ideal theory that will be discussed in Section III: English law learned how to specify rights gradually, by identifying breaches of them.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

Rawls was drawn in a better direction with his emphasis on the basic structure, his understanding of the importance of stability, and his argument about the "strains of commitment" as a limit on what we can do with the full compliance assumption. We can perhaps imagine a more balanced Rawlsian theory that was explicit about the disanalogies between justice and fairness, between macro or structural considerations, and micro or personal ones. And some readers of Rawls (particularly readers of *Political Liberalism)* might reasonably think that the real Rawls offered a theory with that kind of balance, that my argument here is more with Rawlsians who have taken the idea of ideal theory and run with it than with Rawls himself.[14] I acknowledge that this might be right; Rawls is certainly among the thinkers who are often more interesting and complicated than they appear in the hands of their partisans.

But it seems to me that, by according categorical priority to ideal theory, Rawls himself upset the possible balance. He also, I would add, left his theory vulnerable to attacks such as Cohen's that sought to reconcile the contradiction in the wrong direction. Aiming at an ideal theory of *justice* in a general way, where "justice" is understood to be a potential feature of social institutions taken together (and especially if it is their "first virtue"), leaves us without the social friction to which "justice" is a solution. If we could stipulate full compliance with moral rules however demanding, then there is no reason not to stipulate better virtues than justice and a morally good enough humanity not to need a coercive state at all.

It is sometimes thought that a famous argument by Gregory Kavka meets this objection, as it purports to refute Madison's "if men were angels" claim.[15] He argues that even angels would suffer from good-faith moral disagreement (for example, about abortion) and good-faith practical disagreement about the relative weights of moral concerns (for example, about whether second-hand cigarette smoke is sufficiently harmful as to settle the question about whether one has a duty to refrain from smoking around other people) in ways that would sometimes make reliance on individual moral judgment impossible, and would require recourse to collective decision making. Moreover, he argues, we could not assume that angels would reach the needed collective settlements without coercion, as they would *also* have good-faith disagreements about when second-order social moral concerns like "live in peace with your neighbors" and "agree with what has been settled" have priority over their first-order moral disagreements over a question like abortion. He concludes that there is no amount of sheer moral virtue that would lead us to *not* need institutions

[14] For an argument that Rawls's view was more compatible with political realism in the fashion of Bernard Williams as he first appears, see Robert Jubb, "Playing Kant in the Court of King Arthur," *Political Studies* 63, no. 4 (2015): 919–34.

[15] Gregory Kavka, "Why Even Morally Perfect People Would Need Government," *Social Philosophy and Policy* 12, no. 1 (1995): 1–18. See in reply Eric Cave, "Would Pluralist Angels (Really) Need Government?" *Philosophical Studies* 81 (1996): 227–46.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

that coercively settled disputes, even if we refrained from warring over (for instance) ownership of material goods. He acknowledges the possibility that there could be beings who had infinite moral knowledge as well as infinite moral virtue — archangels — who might extricate themselves from these dilemmas; whether they *would* depends on the ultimate shape of morality that they would know (and we don't). As he notes in a footnote, in Rawlsian terms, this is something like saying that even angels would suffer from the burdens of judgment.[16] This sits comfortably alongside Mark Philp's suggestion as part of a defense of realist theory that "[i]f philosophy or morality could be relied on to give us unequivocal and widely legitimate answers to our problems we would not have much need for politics. That we have politics is symptomatic of the need to establish, identify, and enforce rules to govern people's behaviour in the absence of rational consensus on principles."[17]

The first thing to note is that, even if Kavka's argument rescues government from Madison, it does not rescue justice from Hume, and it is misapplied when it is treated as a defense of ideal theorizing about justice. Kavka does not show that the legalistic virtue of *ius* would have to play any part in the angels' dispute settlement. He takes it for granted, freely using the concept of "rights" in describing their moral disputes, but does not examine the assumption.

More importantly, however, even if Kavka's argument is completely successful, it does not at all justify the Rawlsian practice of ideal theory. It instead shows that both strict compliance and consensus are unavailable even in the face of full good faith and good will. It shows that the lack of consensus and compliance are not mere results of *bad* faith and *bad* will. They do not only arise because of Cohen's kidnappers mixing third- and first-person accounts when they demand ransom, or because of Estlund's morally lazy actors who say they "can't" do something when they mean they "won't". In other words, the full compliance assumption that makes clear sense in micro cases does not make sense in the domain of *political* justice. Partial noncompliance at the micro level might be the result of bad will, but partial compliance with political justice is a central feature of political life. This connects the Humean circumstances of justice with what Jeremy Waldron called, in a deliberate echo, the "circumstances of politics," which prominently feature a measure of disagreement about questions that unavoidably require shared answers.[18]

One objection to ideal theory could be: we should not assume full compliance because that assumes comprehensive moral good faith, and if we

[16] It also comes close to Jeremy Waldron's account of the "circumstances of politics"; see Waldron, *Law and Disagreement* (Oxford: Oxford University Press, 1999) — and, as noted above, to ideas developed by Hannah Arendt, whose influence Waldron acknowledges.

[17] Mark Philp, "Realism Without Illusions," 635.

[18] Waldron, *Law and Disagreement*, 102.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

had that, we would have no need of government at all. Kavka's argument, if it is right, defeats *that* claim, by showing the irrelevance of good faith to compliance problems, and by connecting compliance problems directly to dissensus. Here it is worth recalling that a fundamental kind of consensus stands alongside strict compliance in Rawls's idealizing assumptions: he models justice for a well-ordered society, in which "everyone accepts and knows that the others accept the same principles of justice."[19] What Kavka's argument shows is that the two idealizing assumptions are entwined: if we did not have consensus, then we could not assume strict compliance. So long as we are in the circumstances of politics, our theory of (macro, political) justice will have to solve compliance problems. And the circumstances of politics arise even against the background of moral good faith, because of morally innocent disagreement (again, recalling the burdens of judgment). Kavka can be used to defeat the limited-beneficence challenge to Rawlsian ideal theory only by emphasizing especially strongly the circumstances-of-politics challenge. One way or the other, we are left doing political philosophy and theorizing about a political concept of justice precisely in order to solve compliance problems. Kavka identifies a different sort of friction in the system than does Hume, but the result is the same: a theory of political justice is a response to friction, and imagining the friction away would leave us with an enterprise that made no sense.

Much the same is true of an important argument offered by Estlund.[20] Estlund maintains that, Hume notwithstanding, justice could have application — be a relevant category of moral evaluation and obligation — even where there is no need for institutionalized rules or mechanisms of enforcement, even when there are no differences of opinion or conviction, and even if no one is morally deficient in the slightest. "My argument is not," he says, "that justice could have application even if *none* of these things were the case, and that is evidently not true." He means this as an argument in support of more ambitious and demanding theories of justice than Humean accounts might seem to allow. But, like Kavka's argument, it also shows that friction can have more than one source. Good-faith disagreement or a need for rules and institutions can substitute for limited beneficence in the circumstances of justice.

On the one hand, understanding that reasonable disagreement suffices to put us into the circumstances of justice defeats Cohen's effort to allow an extrainstitutional fraternal ethos to swallow justice. This may answer my objection above, that the full compliance assumption raises the possibility of imagining such ambitious moral goods for us to comply with that it removes us from the realm of political philosophy altogether. On the other hand, it all but eliminates the analytical traction that full compliance was supposed to be good for in the first place. Full compliance as a modeling

[19] Rawls, *Theory of Justice*, 4.
[20] David Estlund, "What Is Circumstantial About Justice?" in this volume.

assumption requires consensus. By this I do not mean that people *couldn't* comply with agreed-upon rules even in the absence of a consensus supporting them; indeed, compliance in this circumstance is a normal feature of reasonably stable social and political life. I mean rather that we cannot treat full compliance as an assumption without an accompanying assumption of consensus. (Rawls was right to combine these assumptions.)

If we reasonably disagree about conceptions of justice, and if either we plausibly disagree about the second-order question of when to comply with rules we consider unjust, or we agree on a second-order answer that calls for sometimes not complying with unjust rules (while disagreeing about the first-order question of which rules are unjust), then dissensus entails the possible absence of full compliance. One agent's failure to comply is another's substantive disagreement.

So if we say that the circumstances of justice include either limited beneficence or reasonable disagreement, or both, then Rawlsian ideal theory, or full compliance theory in general, still assumes away the circumstances of justice in modeling justice in action.[21]

This in turn means that modeling partial compliance or disagreement does not taint our thinking about justice in the manner that ideal theorists claim. There might still be times when modeling strict compliance with and consensus around *some* set of norms would be useful, but these would be possible idealizing assumptions among others, not the categorically distinct prerequisites ideal theorists claim. And their usefulness would have to be *defended* like the usefulness of other idealizing assumptions is, not taken for granted. Perhaps sometimes there would be reason to overcome an objection like David Schmidtz's:

> In the real world, the degree of compliance is an endogenous variable, which means that the degree of compliance depends on what people are being asked to comply with. It is not respectable to construct an "ideal" theory while setting aside the problem of compliance, as if compliance issues become relevant only at a later stage of practical implementation. In fact, when we choose a principle of justice, we choose the particular compliance problem that goes with it. If it would take a KGB to achieve compliance, the principle is a bad principle, period.[22]

But the reason to overcome the objection would have to be shown each time. It could not simply be that allowing compliance considerations to

[21] I should emphasize that my remarks on Kavka and Estlund here are conditional: *even if* it is the case that we can replace limited beneficence with disagreement, and so on, I continue to suspect that limited beneficence is crucially relevant to the conditions of the emergence of the norm of justice within human societies, and that those conditions of emergence are more relevant for our continuing understanding of justice than is allowed by Estlund in "What Is Circumstantial About Justice?" But I do not provide any argument for those suspicions here.

[22] David Schmidtz, *Person, Polis, Planet* (Oxford: Oxford University Press, 2008), 8.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

enter amounted to giving a moral veto to those of bad will or bad faith. Without that reason to think that full compliance is a necessary logical feature of a genuine theoretical account of a norm or value (or of justice in particular), the case for treating full compliance as a defining feature of a categorically distinct stage of theorizing that must precede sound thinking about what is to be done fails.

Schmidtz here points to one further asymmetry between restaurant fairness and political justice. Our agreement about the fair division of the dinner bill is capable of transparent enactment and enforcement, just among us, if only we are willing. Political justice is normally taken to be in part a set of norms that are enforceable by an external coercive agency such as a state. And states are not, as I have put it elsewhere, mere machines for dispensing justice.[23] They are social institutions with organizational dynamics and tendencies of their own. Even if the restaurant and political case face a similar structure around the question "will we abide by this rule?" they face a radically different structure when we introduce the question "what will states do, when tasked with enacting and enforcing it?" This is a kind of compliance problem, the kind pointed out by the second half of Madison's dictum: "If angels were to govern men, neither external nor internal controls on government would be necessary." It seems to me a strange feature of the ideal theory literature that it has focused so completely on the question of whether compliance among the citizenry is a valid modeling assumption, to the neglect of the tacit assumption of compliance by the state.

With "full compliance" understood to be a *confined* kind of compliance, operating at micro and mezzo but not macro levels, or on some issues but not all, or within the boundaries of Humean justice, it becomes just one idealizing modeling assumption among others. The same will be true for the strict publicity principle and consensus requirements. They might be useful in particular cases, as means to the end of getting better judgments of practical reason; but they would not be ends in themselves, technical criteria of idealness that must be met for a theory to categorically count as providing normative reasons.[24]

Rawls also adds to the intellectual exercise of ideal theory the need to imagine a society that is unburdened in a variety of ways, including by histories of injustice. We seek to understand the shape of justice, imagining such histories away; we then (it must come sequentially later) measure

[23] Jacob T. Levy, *Rationalism, Pluralism, and Freedom* (Oxford: Oxford University Press, 2014), 58.

[24] In my review of John Tomasi's *Free Market Fairness* (Princeton, NJ: Princeton University Press, 2012), I suggest that Tomasi falls into this trap, being excessively concerned with showing that "free market fairness" counts as an ideal theory at the cost of probabilistic arguments that might leave his theory much more plausible. Jacob T. Levy, review of *Free Market Fairness*, *Journal of Politics* 75, no. 2 (2013): 1–3, DOI: http://dx.doi.org/10.1017/S0022381613000303

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

the real world against that standard and worry about how to respond to the injustices that have already taken place. Arguably, this is the full-compliance requirement projected backward in time: we seek to understand justice as if everyone had *always* complied fully with its requirements, and defer considerations of remedies until after we have identified the norm against which injustices can be measured.

This is, again, initially plausible. In ordinary private law justice, the counterfactual case in which no one had wronged anyone is a necessary baseline. Who would have owned the property in the absence of any theft? What would each of the parties have received if the contract had been performed? How healthy would I have been if you had not negligently injured me, or how much would my goods be worth if you had not vandalized them? It seems sensible to think that this would scale up. In order to know *that* there has been a history of injustice to respond to, in order to know how severe it is and what responses might be called for, we first need to know what the counterfactually just world would have been.

But here again, what is true at the micro level is not true at the macro level. As the magnitude and the duration of the injustice increase, counterfactuals become increasingly fantastic and therefore decreasingly likely to provide a useful baseline. And in political life, those magnitudes and durations have been long indeed. Hume's comment that "Almost all the governments which exist at present, or of which there remains any record in story, have been founded originally, either on usurpation or conquest, or both, without any presence of a fair consent or voluntary subjection of the people" errs only in its first word. When Justice Marshall held that "conquest gives a title which the courts of the conqueror cannot deny," he gestured toward the problem. Once injustice has entered into the foundation of the political and legal order — *and it has always done so* — it becomes implausible to use the no-injustice counterfactual history as a baseline within that order. We could not expect, for example, American courts to adjudicate the legality of the Declaration of Independence and reach the conclusion that the United States' existence was illegal. One need not be a thoroughgoing legal positivist to see that a legal system cannot coherently overthrow its own rule of recognition or *grundnorm.*

But what the legal system cannot do in practice, we are called upon to do in theory if we must construct a no-injustice counterfactual baseline in order to diagnose or propose remedies for historical injustice.[25] We need not even reach the exotic but genuine personal identity problems with pleading for compensation for wrongs using a counterfactual in which none of the living parties would exist because history would be far too different. Reparations for slavery, for example, cannot be grounded on a

[25] My thoughts here have been significantly shaped by many conversations with Timothy Waligore, and by his "Rawls, Self-Respect, and Assurance: How Past Injustice Changes What Publicly Counts as Justice," *Politics, Philosophy, and Economics* 15, no. 1 (2016): 42–66.

claim about what resources some particular extant African-American person would have had in the counterfactual world with no Atlantic slave trade and no American slavery; this person would not exist in that world. The problem is simpler. The supposed ideal-theoretic baseline from which injustice is measured, *isn't.* We cannot imagine away all of our institutions and social facts at the same time. If we must do so in order to move on to questions of what to do in the face of real history, then we will never move on. We do move on regardless, and try to respond to histories of injustice in a way other than the application of the principles we would have. We do in theory as Marshall did, however imperfectly, in legal practice: think about what justice demands of us now, in circumstances we would not have imagined had we not inherited them, the circumstances of a past that cannot be imagined away, made up of injustices that must not be ignored.

III. Epistemic Arguments

The defender of ideal theory might at this point introduce another, subtly distinct, important claim: "[U]ntil the ideal is identified," as Rawls put it (and Simmons agrees) "nonideal theory lacks an objective, an aim, by reference to which its queries can be answered."[26]

This seems to be a logical claim about a relationship between theories — without ideal theory, nonideal theory *lacks an objective* — but it is actually epistemological. "Until the ideal *is identified"* — by us humans who are going to be called upon to judge — "nonideal *theory* lacks an objective, an aim, by reference to which its queries *can be answered"* (italics added) by us humans who are trying to use a theory to answer questions. It is not, on this account, that justice is logically prior to injustice, any more than light is logically prior to dark. It is rather that we who seek to respond to wrongs must begin by identifying what is right.

When we see that this can only be a claim about moral knowledge and learning, however, its intuitive appeal falls away. We know that, for example, people have formulated ideals of freedom in response to the social fact of slavery.[27] They have learned about an ideal by seeing it violated, not the other way around. I think that Adam Smith's model of moral learning in *The Theory of Moral Sentiments* is correct about this: we encounter wrongs that excite our displeasure, then we think about and generalize about what makes those wrongs wrong, and enter a process of moving back and forth between particular cases, often particular cases of wrongs, and general principles, sometimes principles about what is right.

Although Judith Shklar admired Rawls's attempt to theorize in a "realistically utopian" way, this epistemological justification of ideal

[26] Rawls, *Law of Peoples,* 90.

[27] Orlando Patterson, *Freedom in the Making of Western Culture* (New York: Basic Books, 1991).

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

theory stands in plain contradiction to her powerful arguments to "put injustice first." Both Shklar and Iris Marion Young have demonstrated the epistemic power for political theory of accentuating the negative.[28] And once we see that putting injustice first is at least a *possible* epistemic strategy, categorically ideal theory cannot be claimed to have epistemological priority.

If we somehow gain access to moral truth about *either* justice or injustice as a starting point, knowledge about the other category follows. It is no more absurd to begin by identifying a direction to move away from than to begin by identifying the direction of the ultimate goal.[29] We have no reason to treat the light as *logically* prior to the dark; the supposed advantages of ideal theory are *only* epistemic: it is easier to *see* the light than the dark. With respect to normative questions, it seems to me that this is, in general, false. Outrage against and punishment of injustices are closer to being basic facts of our moral psychology than are the articulation and abstraction of principles of justice. Individually and collectively we (at least often) begin our moral learning with negative cases, not (or at least not always) with positive theories. Rawls was ready to "assume that insight can be gained in no other way," and that assumption made categorically ideal theory necessary. But it is simply false, and its falseness makes unavailable this epistemological argument for ideal theory.

In my account of Smithian moral learning as involving a back and forth process between cases and principles, evils and ideals, the reader sympathetic to Rawls will of course hear an echo of reflective equilibrium. Such tacking back and forth is so basic to our moral psychology that it would be surprising *not* to find something like it in many different theories. But reflective equilibrium need not begin with idealizing assumptions; it could begin with sustained reflection on evils and harms, then proceed to thinking in a general way about those evils and harms, looking for principles on the injustice side. I think both Shklar and Young do work like this. Nor must reflective equilibrium necessarily issue in a settled set of ideal principles. This marks an important limit to the echo. The Rawlsian sequencing of ideal and nonideal *limits* the tacking back and forth, reflective equilibrium notwithstanding. Once we have formulated an ideal theory of justice — which certainly involves the use of reflective equilibrium and allows for the consideration of some grave evils as fixed points of our moral judgment — we set about to apply it as best as it can be applied in

[28] Shklar, *The Faces of Injustice;* Iris Marion Young, *Justice and the Politics of Difference* (Princeton, NJ: Princeton University Press 1990). For an account of intellectual progress in normative institutional analysis based on avoiding evils rather than on ideal theory, see David Wiens, "Prescribing Institutions Without Ideal Theory," *The Journal of Political Philosophy* 20, no. 1 (2012): 45–70.

[29] While I do not agree with all of Sen's arguments against ideal theory in *The Idea of Justice* (Cambridge, MA: Harvard University Press, 2009), the epistemic argument represented by the mountain-climbing metaphor — we do not have to know the location of Mount Everest in order to know the direction of "uphill" — seems absolutely correct.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

nonideal circumstances. We don't, however, learn from that process in a way that could be legitimately imported back into the now-settled ideal stage of theorizing. To do so would be to contaminate the ideal theory.

On this, in their various intellectual idioms, I think Cohen and Estlund agree with Rawls and Simmons. For Cohen, the distinction between what we should think about justice and what is to be done in the world is a hard one, and the former is not informed by the facts that we learn by trying to do what is to be done. For Estlund, we must preserve the space for being able to think about justice in ways that lead to ambitious and demanding conclusions. Allowing, for example, failures of implementation in the world to feed back into our understanding of what we should morally aim at is to provide an opportunity for weakness of will, bad faith, recalcitrance, or wickedness to entrench themselves, and to deprive ourselves of the critical standpoint with which to describe injustice correctly. Although Cohen does not share the constructivist project of finding a procedure for selecting principles of justice, and Estlund is not committed to it, they both have reason to endorse this aspect of the constructivist procedure.

One further epistemic claim is made on behalf of ideal theory: that its willingness to identify far-off and ultimate goals can protect against myopia, short-term thinking, and merely local normative optima. We do not need to know the location of Everest to identify uphill and downhill *on this hill*, but our goal should not be to climb whatever hill we happen to stand on. Simmons emphasizes, along with full compliance, the function of ideal theory in keeping our attention on possible paths to higher hills, in preventing us from being trapped in second-best equilibria. Only with ideal theory in hand can we identify the path we should be taking.

Here there is considerable risk of subsuming the supposed epistemic and methodological case for ideal theory within a substantive normative position. The (in my view extremely dubious) advice to let the perfect always be the enemy of the good, to prefer the formal possibility of attaining the best to the probability of attaining the better, does not depend on the ideal theoretic strategy for *identifying* the best. We might put injustice analytically first and still follow this rule about what is to be done. How we should weight short- and medium-term improvements against faint possibilities of long-term triumphs is an entirely separate question from whether we can identify or imaginably reach those triumphs. Moreover, the theorem of the second-best shows that there is no necessary relationship between hitting some far-off normative target and moving in its direction. While the theorem of the second-best is compatible with ideal-theoretic identification of the first-best, it is *not* wholly compatible with the view that the value in such identification is that it gives us a normative target to work toward, or that it protects us against getting caught in some locally good result at the expense of the pursuit of the perfect.

A full compliance assumption does, as we have seen, have genuine epistemic benefits for local and immediate micro-scale fairness, even if I am

right that it does not scale up. But the case for epistemic advantages in these long-term cases is much less clear. I think the argument is tacitly Kantian: because we *must* be able and willing to preserve the possibility of progress, we must also assume that we can identify the path toward it. But this is methodologically suspect at best. A moral imperative to be able to identify a path does not generate the actual ability to do so in any reliable fashion.

This emphasis on the far-off ideal also conflates the method of ideal theory with the distinct substantive merits of normatively *ambitious* theory. Until now none of my arguments about ideal theory have been addressed to its ambition as such, and deliberately so. The idealists are right to worry about leaning too much on constraints of *political* feasibility in particular. (This is the germ of truth in Estlund's "utopophobia" argument.) Some opponents of ideal theory have stressed the need for attainable reformist goals, and denied that ideal theory as traditionally conceived can provide them; if political philosophy is to be action-guiding it must offer us things to do that are doable. A normative vision that is too far distant will be met with the traditional New Hampshire shrug on being asked for directions: "Y'can't get theah from heah." Normative theory should, so this line of argument goes, show us places to which we *can* get from heah.

I am temperamentally a reformist, not a revolutionary, but I recognize that this line of argument is easily overstated. The idealists' responses to it as a general proposition are correct. Political philosophy is also not meant to tell us to stand still, and insisting on only articulating goals toward which the path is obvious and smooth offers opponents an unacceptable level of authority over our normative aspirations. It cannot be the case that my otherwise-right ideal becomes wrong simply because people whose moral views are wrong decide to stand in its way. It may become harder to reach; it may even become impossible to reach; the right thing for me to do, all things considered, in the face of that may be to compromise. But my opponent in a normative debate is argumentatively estopped from treating the fact of his own objection as proof of my wrongness. "You say that justice demands equality, but I forthrightly oppose equality and I have enough power to stop you from implementing equality, therefore your goal is not attainable, *therefore you are wrong about what justice demands;"* this does not and cannot follow, and we have to be able to name compromise as compromise.[30] Constraints of economic and sociological feasibility are usually *not* like this, but constraints of strictly political feasibility *sometimes* are — and I do not think that normative theory is limited to describing goals that do not happen to contradict the desires of the powerful.

[30] Compare G. A. Cohen's discussion of the talented holding society to ransom for the use of their talents in *Rescuing Justice and Equality,* a discussion that I think is substantively entirely wrong but structurally very illuminating.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

I do think there is such a thing as normative theorizing that is so ambitious as to be pernicious.[31] I am, as Estlund would have it, a utopophobe, thinking that the embrace of distant and barely-conceivably-realizable goods is often a distraction from the possibilities of plausible improvements, that dreams of omelets encourage an unhealthy eagerness to get on with breaking eggs. But those psychological and sociological indictments of the *practice* of ambitious theorizing are not the point of this essay. Estlund is right that concerns such as these are distinguishable from the question of the conceptual or logical standing of utopian thinking.[32] My objections are not to the ambitious conclusions of ideal theorizing, but to its mistaken beginnings; its conceptual standing is precisely what is at stake.

Simmons's emphasis on the importance of the distant moral horizon conflates the two. It treats the value of very ambitious theory as a reason to insist on the priority of ideal theory. This, it seems to me, is simply a mistake. Moreover, methodologically ideal theory is neither necessary nor sufficient to generate utopian or perniciously ambitious theory, even if there is a kind of elective affinity between them and even if they appeal to the same kind of temperament. By my lights, Marx counts as a perniciously ambitious theorist, but he needed nothing like the methods of ideal theory to get him there. Putting injustice first and bearing witness to the evils of the world can generate the immoderate desire to sweep away *all* injustice and evil, costs be damned. The method of observing evils first is compatible with finding that the whole world is so filled with them as to call for truly radical change.[33]

Here I should add that methodologically ideal theory is neither necessary nor sufficient to generate *non*-perniciously ambitious theory, either. This is one

[31] For reasons that overlap in part with those of Charles Mills's important article "'Ideal Theory' as Ideology," *Hypatia* 20, no. 3 (2005): 165–83. It seems to me that the force of Mills's argument remains widely underappreciated. Pointing out that ideal theory can yield very ambitious and demanding normative reform agendas is not, or at least not straightforwardly, responsive to his critique. The presentation of ideal theory as an idealized abstraction from real states of affairs — real democratic processes, real constitutional histories — makes it difficult to diagnose those processes and states of affairs as themselves primary sources of injustice. Rawls freely acknowledges, in *Lectures on the History of Political Philosophy* (Cambridge, MA: Harvard University Press, 2008), that his conception of political philosophy includes a reconciliatory feature; it helps citizens to see the moral value in what really exists and the possibilities for moral improvement that lie immanent within it. Political philosophy can serve to legitimize real institutions in order to make moral reform within them seem attainable. But — I take Mills's point to be in part — that kind of legitimation and reconciliation can also serve to mask existing institutions' deep complicity with or commitment to injustice.

[32] This implies that the Stanley Fish reference in the title could be completed: there's no such thing as ideal theory and it's a good thing, too. But the goodness of the thing is not the point of this essay, and I mean my argument to be persuasive even to those who would regret the realization that ideal theory is unavailable. By contrast, see Colin Farrelly, "Justice in Ideal Theory: A Refutation," *Political Studies* 55, no. 4 (2007): 844–64.

[33] See also Lorna Finalyson, "*With Radicals Like These, Who Needs Conservatives?* Doom, Gloom, and Realism in Political Theory," *European Journal of Political Theory*, forthcoming, DOI 10.1177/1474885114568815.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

of the moments in which the conflations among idealism, idealization, and ideal theory have come to serve us badly. The abolition of slavery is sometimes cited by ideal theorists as a crucial case for their rejection of reformism in general and for their refusal to allow recalcitrant opposition to feed back into the formulation of ideals in particular. The abolitionists were radicals committed to the truth that slavery was morally intolerable, and appropriately kept their eye on the distant moral horizon of abolition. This is all true; but it does not make the abolitionists into ideal theorists. They put injustice first if anyone did. While there were utopians among them, and those who spun elaborately complete theories of liberty, abolitionism as such did not depend on such theories. The distant moral horizon was ambitious and difficult; but it was set by evils. There was no need, perceived or real, to settle on a fully worked out theory of liberty before appreciating, and rejecting, the evil of slavery.

IV. Justice and Politics

I mean to suggest that the Humean circumstances of justice, or something close to them, are not waivable assumptions in theorizing about justice; they help to set the meaning of the inquiry. Indeed I do not think the Humean circumstances of justice exhaust the list of non-waivable considerations. Kant's "unsocial sociability" is a considerably broader concept than Hume's limited beneficence, but is no more expendable. Humans are naturally sociable creatures who cluster in groups smaller than humanity as a whole with contentious boundaries between them; naturally cooperative and competitive at the same time; working together in groups but comparing themselves individually to their fellows in moments of *amour-propre;* prone to live by norms conducive to group cohesion and to harshly punish offenses against those norms. Mutual vulnerability (Hobbes), mortality (Heidegger), natality and plurality (Arendt), limited knowledge and information (Hayek, Oakeshott), the circumstances of politics (Waldron), linguistic communication rather than telepathy — these are the kinds of considerations that enter in at the beginning of theorizing the human, social, political condition. Perhaps some of them are more fundamental than others; perhaps some are not fundamental at all. (I have never understood Arendt's view that the human condition is partly constituted by being terrestrial and would be compromised or transformed in some deep sense by long-term space travel.) But it is with good reason that the great theorists so often began with their various accounts of human limitations, and a theory that is constructed without reference to *any* of these is neither a theory of justice nor a theory of politics but a branch of speculative fiction. To use the unpleasant old joke: once you're theorizing about *justice,* once you're doing *political* philosophy, we know what kind of theorist you are, and we're just haggling about price.

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

https://doi.org/10.1017/S026505251600025X Published online by Cambridge University Press

There is a legitimate wide range in how much one idealizes in normative theory, a spectrum from taking the world as fixed to taking it as very susceptible to reform. But it is a spectrum, not a ranked, categorical distinction. And so I do not mean my critique of ideal theory to count as a defense of what is now thought of as nonideal theory: theory that takes what ideal theory gives it and then works to solve questions of compromise or attainment or rectification. My conclusions more resemble the political realist critique by Williams and his followers than current nonideal theory done in ideal theory's shadow. The principles appropriate to political and legal life are partly *constituted* by the problems and limitations of human social life; they are not imported from the realm of moral truth and then applied to a more or less recalcitrant world.[34] And if we are always engaged in some normative imagination that the world is other than it might be, it is still *the world* that we are imagining: this world, not some other one to which our sense of norms might not apply especially well.[35] I quite like speculative fiction, and have spent more than one late night discussing a topic like "how would society be organized if we had technology that succeeded in defeating the problem of scarcity, or immortality, or a *Foundation*-style solution to information problems?" This kind of thing is good fun and can sometimes cast interesting light on some feature of social organization. That does not make it something we *must* engage in, or even usually *should* engage in, as a first step to understanding the question of practical reason, what is to be done in the world we inhabit.

Political Science, McGill University

[34] I discuss this idea in greater depth, and with reference to Augustine's distinction between the principles and ethics of political life in earthly Babylon on the one hand and the moral truth available in Heaven on the other, in Jacob T. Levy, "Against Fraternity: Democracy Without Solidarity," in Will Kymlicka and Keith Banting, eds., *The Strains of Commitment: The Political Sources of Solidarity in Diverse Societies* (Oxford: Oxford University Press, forthcoming).

[35] It would be customary to say "our normative intuitions," but I think that word understates the likelihood that our normative sense is an evolutionary feature that — at least at relevant scales — provides reliable information about moral truth or what is good for humans in society. As social conditions or the circumstances of human social life change, they become less reliable guides. I don't think we can actually know much about morality within a race of invulnerable immortals, for example. See also Jakob Elster, "How Outlandish Can Imaginary Cases Be?" *Journal of Applied Philosophy* 28, no. 3 (2011): 241–58.