

## Real World Theory, Complacency, and Aspiration

Geoffrey Brennan                      and                      Geoffrey Sayre-McCord  
Philosophy Department ANU/UNC      Philosophy Department UNC-Chapel Hill

### I Introduction

In *Utopophobia*, Estlund describes his project as a methodological one. Its object, he tells us, is to put “*idealism and realism in their proper places*” in political philosophy (p 4). His discussion focuses (as shall we) on justice — but the arguments are more general and the issues they raise reach across all of political philosophy and more broadly. They concern which lines of enquiry and what insights are relevant to a theory of the appropriate criteria for the evaluation of our political institutions and social interactions.

The issue for Estlund is “*whether theories of such things as justice are incorrect or are in some way a failure if there is no expectation of justice as understood by the theory being achieved.*” (p4) He denies that theories that have this feature are incorrect or a failure — or more accurately, he denies that anyone has shown decisively that they are.<sup>1</sup> Yet he worries that much theorising about justice effectively “*forecloses ... the possibility that human societies might be prone (perhaps indefinitely) to be significantly unjust*”; and moreover, forecloses this possibility merely “*by assumption or definition*” (p4).

Against this, Estlund sets out to defend the idea that a theory of justice which “*requires more than we can ever expect to achieve is not thereby in the least flawed*” (p4). Only if this is admitted, he maintains, will there be room in political philosophy for an approach that is “*non-complacent and non-concessive*” (p6). The complacency issue is one that we shall take up in more detail in section IV below, but we will set it aside for now. The concern about concessiveness is, conceptually, more central to the logic of Estlund’s project (though not perhaps to the psychology), and we turn to this issue immediately.

A theory is non-concessive, in the appropriate sense, when, in developing it, constraints that arise by virtue of any crookedness in the timber of humanity are set aside. In particular, non-concessive theories assume that, whatever the principles of justice are, those principles will enjoy full compliance by those subject to them. This of course involves idealizing away what is an all too familiar feature of human beings: that they sometimes fail to do what they themselves accept as right. Following what has become standard practice, in what follows we will refer to such utopian theorizing as an instance of “*ideal theory.*”<sup>2</sup>

---

<sup>1</sup> Not only is Estlund’s main argument a negative one – that criticisms of so-called Ideal Theory are not decisive -- it is also notable that the two central attributes in the approach Estlund hopes to defend – non-complacency and non-concessiveness -- are expressed by him in negative form.

<sup>2</sup> Perhaps it is worth noting that along with certain idealizing assumptions, ideal theory also regularly takes as given, various features of the human situation that are, pretty much by all accounts, seen as less than ideal. What leverages a theory into the category of ideal theory is not that it idealizes away from all things that we might wish were different, but that it takes certain features that we recognize are almost surely not going to be realized and nonetheless build them in to the conditions under which we judge the acceptability of candidate principles of justice.

To be clear, Estlund is not against concessive reasoning. “*Much thought about social justice is in a concessive mode, taking people ‘as they are’ ... and so forth. There is*”, he remarks, “*nothing wrong with that, of course.*” (p6) However, in Estlund’s view, a ‘non-concessive’ approach to political philosophy is more fundamental conceptually. That approach engages the correctness or truth about justice in a way that he thinks a concessive approach does not. He concedes that one upshot of his view is that it may well be that “*no-one should ‘set out for justice’*”; after all, as he acknowledges, “*futility is plausibly a defect in any practical project.*” Estlund’s response to this consideration is to draw a distinction between the *correctness* of an understanding of justice and its practical value, claiming that “*the connection between sound political philosophy and guidance for political practice has often been exaggerated*” (p4).

While we grant that, very likely, the connections have often been exaggerated, we think that there are important connections. Moreover, as we see things, those connections, contra Estlund, call for keeping a close eye on what people can and will do when evaluating candidate principles of justice. This is what we mean by “real-world theory”. Indeed, we suspect an utterly “non-concessive” theory of justice is likely to encourage demoralization and/or unjustified recklessness in contexts where justice’s true demands stand the chance of inspiring and successfully guiding agents concerned with justice. (This is a possibility we explore in what follows.) At the same time, we take seriously Estlund’s caution against being (too) concessive and (unacceptably) complacent. Just how these concerns can be squared with a view that sees theories of justice as offering practical, not just theoretical, criteria will be our main focus in this paper.

In what follows, we shall focus our attention on four aspects of Estlund’s argument. The first (section II) involves Estlund’s strategy of assigning the onus of proof to those he defines as his opponents. The second (section III) explores Estlund’s central distinction between concessive and non-concessive approaches to political philosophy. We think that that distinction rests on an untenable view about “where to draw the line” between the acceptable and the unacceptable. What makes it untenable, we will argue, is that it fails to accommodate properly the ineliminable need for agents (and institutions) to make trade-offs among considerations countenanced as relevant by, and to, justice. The third (section IV) involves the ‘complacency’ charge Estlund levels at ‘concessive’ approaches to political philosophy. The fourth (section V) addresses the role that an understanding of justice, in its non-concessive form, plays as a method of understanding the concept. The central issue in this last section is a claim of John Rawls’, which Estlund quotes, but is cagey about endorsing, to the effect that ideal theorising of the kind Estlund does endorse is a necessary precondition for a proper understanding of justice and hence for the associated practical project of its implementation. Predictably, we will push against Rawls’ claim.

## **II The Strategy of Argument**

Estlund acknowledges the value of the ‘concessive approach’ — of ‘non-ideal theory’ as it is more commonly called (‘real-world theory’ as we prefer to call it). What he denies is that the value of non-ideal theory tells against the value, correctness, or legitimacy, of ideal theory. He characterizes those who maintain that in doing political philosophy we should take account of what is feasible given human failings, as “utopophobes” on the grounds that they sweep ideal theory off the table without adequate reason.

One might think that that makes Estlund a “utopophilic” — but this would be a mistake, for two reasons. One is that Estlund makes it clear that he entirely rejects utopian ‘blueprints’ (as too specific); and indeed he rejects ‘utopian proposals’ (on ought-implies-can grounds). But

‘utopian *principles*’, he insists, are, so far as anyone has argued, perfectly acceptable. And exploring the content of those principles is correspondingly a perfectly acceptable intellectual project. Accordingly, unless those who think otherwise can provide compelling reasons for their view then they should allow intellectual space for those who prefer to focus on ‘ideal theory’ — the non-concessive approach — to pursue their enquiries uninhibited.

In effect, Estlund is issuing a “put up or shut up” challenge to the “utopophobes”. That challenge is to produce a decisive argument as to why the ‘non-concessive’ approach should be rejected. The onus of proof then lies with the ‘real-world theorists’. Estlund does more than locate an onus of proof of course. He offers a range of arguments for thinking the non-concessive approach has conceptual value. But even if those arguments were inadequate or unpersuasive, the ‘onus of proof’ challenge to the utopophobes would still stand.

We like to think of ourselves as dispositionally tolerant souls. And to the extent that Estlund’s plea is that, in the final analysis, people thinking about justice have a perfect right to pursue questions they consider interesting and valuable, we have no quibble. Indeed, one might bolster that appeal to rights by observing that the enterprise of understanding justice will no doubt go better when people are free in this way. But Estlund’s charge against utopophobes cannot be that they are restricting intellectual freedom. Rather, it is that they are making a mistake about the value, correctness, or legitimacy of criteria of justice, as criteria of justice, that do not take account of human weakness and failure.

At the risk of being too concessive, we acknowledge that *decisive* arguments, when it comes to how to think about the principles of justice, are exceedingly rare, and not something that we see ourselves as in a position to offer. Straightaway, then, we grant that those who reject Ideal Theory have not given decisive arguments on their side. What there is instead is a ‘weight of considerations’ for and against various views. And in that setting, we find ourselves on the side of real-world theory, though not, we think it is fair to say, on the basis of any sort of a phobia.

Simply put, we think that the balance of considerations weighs in favour of the ‘concessive’ alternative – and while we agree that Estlund’s core arguments raise concerns, we do not think that they tell against real-world theory properly pursued. As a result, we remain what Estlund would characterize as utopophobes, albeit cautiously and, we hope, with open minds.

### III The Non-Concessive Principle in Principle

Left to itself, the non-concessive approach to justice invites a bi-modal evaluation of the world and the choices we might face in it: either the world, and our choices satisfy the privileged “ideal” principles of justice or they do not, and if they do not, they are, in the relevant sense, unjust.

One can see immediately why a thoroughly ‘non-concessive approach’ is going to be hopelessly uninformative for those – all of us -- faced with circumstances in which full justice cannot be achieved. Strictly speaking, one cannot judge whether either of two situations is more or less just, except if (at least) one of those situations happens to meet the standards set by the ideal theory. Relatedly, the idea of ‘improvements in justice,’ large or small, is not one the approach encompasses. To the extent we are engaged in the ‘practical project’ of securing greater justice, we need a *metric* of justice. And in probabilistic cases, where the outcomes of actions or policies or institutions are uncertain, we need a metric that allows a measure of how just (or how unjust) various possible situations are or would be.

In this spirit, John Broome (1999) asserts that “goodness is fully reducible to betterness”. We are not entirely sure what “full reducibility” amounts to — but Broome’s discussion is aimed to make salient a distinction between a measure of the *degree* of ‘goodness’ (what he calls ‘betterness’) on the one hand, and where along the relevant spectrum from worst to best one should ‘draw the line’ between good and bad, on the other. And this distinction is the one we seek to appeal to here.

We suspect Estlund would be more than willing to grant the importance of shifting from ideal to real-world theory when trying to establish the relevant metrics and when thinking through questions of what justice demands of us, in the circumstances we actually face. But he thinks that, with that point acknowledged, there is still important room in our thinking about such questions for the principles that would be the *right* principles of justice *if* people were free from human failings (such as weakness of will, unjustified self-interest, and moral indifference).

Specifically, as we understand Estlund, the role for ideal principles is to provide a bulwark against lowering moral expectations in a way that would count what we should kick against as intolerable as morally “acceptable”. His worry is that if our standards already take account of our failings, they will, in that concession to human nature, fail to set the standards that are actually appropriate to us as moral agents. This possible worry involves ‘concession’ in an obvious sense. But our commitment to the importance of judgments of relative justice (whether an articulate metric is available or not) does not entail drawing any ‘lines’ between what is acceptable and what is not. Our concern that criteria of justice be able to make comparative judgments of relative justice is independent of where along the resulting continuum the line between what is acceptable and what not is set. Certainly, one issue is whether the line (if one is to be drawn) can be “high enough” to avoid complacency and to engage well-grounded aspiration. There is of course a coordinate question about how to draw the line between an appropriate aspiration, on the one hand, and a dream or fantasy, on the other hand. Even as we grant the importance of aspiration, we are inclined to think standards of justice that fit ideal circumstances, absent a defense of their appropriateness to other circumstances, lack a compelling general claim to being an appropriate target of action.

Estlund emphasizes, rightly in our view, that, however much the world improves from the point of view of justice, we should be able to entertain the possibility that it could be better.<sup>3</sup> Worries about our being unable to entertain that possibility is a key element in his concerns about complacency (which we will turn to in the next section). But he seems to think that the possibility of an improvement can only be entertained if we have the ideal as a benchmark. That is clearly not the case. A metric of betterness (specifically of “more just,” no matter how imprecise) can do that work, if it does not impose a limit, or at least does not impose a limit that we will likely reach. What a measure of betterness, *vis a vis* justice, needs to do is leave room for improvement for as long as improvement is a justifiable aspiration. At the same time, recognizing that things might get *better*, from the point of view of justice, makes salient the concern that things might, sadly, get much *worse* -- though in different ways, and not all of them equally bad. An ideal theory, taken alone, offers no standards for thinking about, and responding to, these latter possibilities. Moreover, additional argument would need to be offered before it would be plausible to claim that thinking well about these depressing possibilities requires an appeal to an Ideal Theory. There is no reason that we see, to think that local improvements, or

---

<sup>3</sup> Of course, absent some way of comparing failures to meet the ideal, and Ideal theory will lack the resources for distinguishing among opportunities except as they realize, or fail to realize, the ideal. When we are thinking about ways our world might actually be, a reasonable suspicion is that every available option will end up registering simply as a failure.

defenses against local decays, are fruitfully thought of as successive approximations to extreme ideals (or for that matter, moves away from anti-ideals).

Estlund might respond that there is good reason to think that we cannot formulate a suitable metric of justice without appeal to the ideal case. Metrics of justice, he might claim, can only be formulated once the ideal has been specified. As we are thinking of this idea, we see it as linking directly to Rawls' claim that there can be no systematic theory of justice that does not appeal to an ideal theory of justice. We examine this claim in detail in section V, and we will leave further remark on it until then. Suffice it to say here that we think that this claim is mistaken.

However, a metric of the kind we have in mind is normative in the sense not just that it distinguishes among levels of accomplishment but that, in the process, it attaches to each such level a measure of desirability. And in this connection, there is no presumption that desirability cannot map into performance discontinuously. There may be a level of accomplishment such that performance below that level is considerably less desirable than performance above that level. Still, and importantly, there are comparisons to be made, inevitably, and a theory consisting of principles suited, so far as the arguments go, only to ideal circumstances are not good measures of the relative value of options in other circumstances. Importantly, Estlund himself appeals not simply to judgments about what meets, and what falls short of, the principles of an Ideal Theory. For example, he offers an illustration of a point by appeal to cases of extreme injustice, which he recognizes to be very bad (p. 23). That judgment requires a metric of injustice, which is something that the concessive/non-concessive distinction does not directly accommodate. At the very least, concessive or not, such examples cry out for judgments both about ways that we might hope the situation could be improved and worries about how it might get yet worse. Any theory that fails to have the resources to address such concerns is, we think, unacceptable. And, at present, we see no reason to think that an ideal theory helps with these tasks – but we do see reason to think that it may well hurt (though more about this soon).

It is worth emphasizing that a non-concessive posture blocks not just concessions to the realities of human nature but also concessions to other legitimate (but less than ideal) norms that might well be in play. Cohen, whom Estlund counts as a fellow ideal-theorist, is not shy about conceding that justice is not the only value; and that one needs some kind of 'value-indifference curves' to make the appropriate trade-offs between justice and other values.<sup>4</sup> Importantly, in order for the trade-offs to be a conceptual possibility a metric of each of the relevant values is required, and not simply a standard that works for ideal conditions. One needs to know how much justice one is giving up in order to secure a given amount of additional equality (or freedom, or aggregate well-being or whatever other value is in play). A 'non-concessional approach' is equipped only with the ideal requirements in each case. That approach can make salient "impossibility theorems" concerning different desiderata, but it cannot resolve them. Cohen, at least, thinks they can be resolved — and indeed resolved as it were, in terms of better or worse, as his appeal to the indifference curve formulation suggests. The necessity for such trade-offs might be a result of pure logic (as in Arrow's famous impossibility results) or upshots of human motivational imperfections. Either way, the non-concessional approach is inhospitable to formulating normative concepts in the appropriate way — a way that focuses not on what requirements ideal justice would require, but on what would make for more or less justice.

Just to exemplify the challenge — here focusing on forced choices between justice and equality: an important step, as Sen is surely right to highlight, is "equality of what?". However, once that issue is settled, we must ask the concessively-hospitable question: which measure of (the

---

<sup>4</sup> See, for example, the final section of Cohen (2003)

relevant) inequality shall we deploy? The variance of the distribution? The Gini coefficient? The number of inequalities? The number of inequalities greater than a specified level? And will the same measure be applicable in different domains? The general question – which measure – is clearly a normative question. Yet it can only arise if the focus of attention is relative assessment within the domain of all outcomes — ideal and real-world alike. Such relative assessment is an aspect to which a concessive approach is not just hospitable, but essential. In contrast, it is a matter on which the non-concessive approach, taken by itself, is silent.

#### IV Complacency, Caution and Realism

When Estlund first introduces the notion of complacency, it is in the context of a form of “*strict realism*” which he thinks “*no one will insist on*” (p5). This form of realism is one where any criticism of the status quo, whatever it happens to be, is ruled out because (or to the extent that) that status quo can be shown to emerge as a unique<sup>5</sup> equilibrium in some kind of social process – and one in which the motives of agents are taken as given. Estlund labels that position “*complacent realism*”: the label is explicitly “*derogatory*” (p5). We certainly do not want to insist on this sort of “strict realism.”

Yet we are moved to register several points about what is at stake in this unattractive position.

First, it assumes that where there is no real prospect of change, criticism is ruled out. With Estlund, we think this misunderstands the conditions of legitimate criticism. Brennan, an Australian, can feel outraged at remarks of Donald Trump’s or of some Brexiteers, and judge them unjust, or offensive, or whatever, even though he is impotent to change what those people have said (or are likely to say or do in the future). Criticism is surely not silenced by an incapacity to change what one criticizes. Indeed, we are inclined to judge someone morally deficient if they fail to condemn things, just because they cannot change those things<sup>6</sup>. Still, that is compatible, to say the least, with thinking that legitimate criticism needs to be sensitive not just to what one can achieve oneself but to what might possibly be achieved (in the way real-world theory supposes).

Second, Estlund’s promotion of a “non-complacent and non-concessive” position on justice carries the suggestion that the concessive alternative actually *induces* complacency: and the avoidance of such complacency thus becomes a *reason* for the non-concessive approach. We are unsure whether Estlund’s reference to ‘complacency’ is supposed to be a term of art characterizing an approach to normative theory or a claim about the psychological impact of that approach. There is a danger here of a kind of rhetorical drift – and we wonder how far that is intended or whether Estlund is a victim of the rhetorical force of his own terminology. Certainly, Estlund believes that real-world theory underplays the important aspirational role of thinking about justice (and for that matter, about other issues in political philosophy). We suspect that the ‘non-complacency’ theme of Estlund’s approach does (and is intended to do) considerable persuasive work. And we are on board. After all, complacency (and the moral smugness it connotes) is repugnant. If specific methods of analysis encourage repugnant attitudes, then that is certainly grounds to worry about those methods.

But there are lots of questions here. What exactly do we mean by ‘complacency’? When might it apply? What is the status of an argument about methods and the determination of concepts that

<sup>5</sup> If the equilibrium in question is not unique then there is presumably an issue of equilibrium choice that would permit some action-guiding element.

<sup>6</sup> The general point is that normative assessment is first a matter of attitudes; and attitudes can be appropriate even though no action by anyone could change things.

appeals to their psychological upshots (as the anti-complacency argument seemingly does)? We want to say a little in this section about each of these matters in turn.

A. *Getting a Fix on Complacency:*

Complacency might be understood either attitudinally or behaviorally – either as a reflexive attitude towards one’s own qualities (in this case moral qualities) or as a disposition to inaction. We should distinguish these senses. So, take the attitude first.

1. Complacency as Attitude:

Suppose A does something admirable. It would be appropriate (other things equal) that any observer B or C admire A’s action, and admire A herself for intending and successfully performing it. Those observer attitudes would be, in the familiar lingo, “fitting”. It would be no less “fitting” for A herself to have that same set of attitudes towards D who had done a similarly admirable thing. But when it comes to A’s attitudes to her *own* action/motives/execution, such attitudes (though entirely in line with those of Adam Smith’s “impartial spectator”) would not, most people feel, be fully appropriate *for A*. She should resist self-admiration and be especially alert to ways in which she might have done yet better – she should have the (additional) virtue of moral modesty or at least not suffer a too easy self-satisfaction. And her failure to have that modesty – her failure to have less approval for her own actions than it would be appropriate for her to have for the actions of others – would likely be widely regarded as a moral failure.<sup>7</sup>

Of course, people in general may be prone to exaggerate their own moral qualities – to think better of themselves than they deserve. They may hold attitudes towards their own conduct that do not satisfy the “fittingness” test: their attitudes toward the equally meritorious actions of others may indeed often be less generous than those they hold towards their own. (Much in the same way as the overwhelming majority of people consider themselves to be, as drivers, “better than average”).<sup>8</sup> We recognize this not just as a possibility but as a rather robust aspect of human nature.

Attitudinal complacency covers both cases. That is, it applies whether or not the attitude towards one’s own moral qualities is “fitting”. It is a meta-level reflexive attitude that can operate independently of the fittingness of one’s substantive attitudes – though of course it is more objectionable the more biased towards the self the self-appraisal is.

It is important to underline the reflexivity of ‘complacency’. One would not normally refer to someone being complacent about *another’s* moral qualities – unless she had reason to think that those moral qualities had been brought about by actions of her own. This leads us to think that if the methods of political philosophy lead to excessive ‘complacency’, the excess is attributable less to whether the approach is ‘ideal’ or ‘real-world’ and more to the failure to distinguish adequately between self-regard and the regard of others (that is, first and third person regard). It seems strained, to say the least, that someone counts as exhibiting attitudinal complacency in cases where her attitude concerns something she knows she had no hand in constructing or supporting. Whether, nonetheless, there is a different sense of complacency that might accurately characterize people, perhaps considered as a group, that is independent of their

---

<sup>7</sup> We are not here endorsing this view, just noting the intelligibility, and general attractiveness, or distinguishing the between the attitudes one might appropriately have to others and those one might have to one’s self.

<sup>8</sup> There is a substantial literature in psychology about biased self-appraisal – or “positive illusions” as they are sometimes called. See Taylor and Brown (1988)

individual actions and accomplishments we will consider later -- but we expect that sense to be similar in structure with complacency about one's own character and accomplishments

## 2. Complacency as Inaction:

'Complacency' might however refer not to a *reflexive attitude* but rather to the disposition to inaction. In this case, and specifically shorn of repugnant attitudes, more seems to be required to establish the force of the critique. While complacency as smugness may be uniformly reprehensible, there does not seem to be the same systematic case for action over inaction. At the very least, there may be cases where doing 'nothing' is better than doing 'something' – including cases where striving specifically for the ideal is worse than alternatives.

Part of the Estlund plea for ideal theory is that we should be alert to the (many) ways in which the world might be better (including more just) than it is. And we agree. But such alertness is symmetric: we should also be alert to the (many) ways in which the world could be *worse* (including less just) than it is. There should be no general presumption that a change to the status quo is more likely to do good than to do harm – or that the amount of good on offer will reliably be greater than the extent of harm that the action risks doing. In lots of cases, the effects of one's action will not be clear: and this may well be especially true of more radical schemes associated with idealist projects. Focusing on the ideal may lead to taking excessive risks. Often it makes sense to take a clear-eyed view of what the various outcomes of an action might be, to assess those outcomes as accurately as one can. Sometimes, too, that assessment leads to discovering that inaction actually promises more justice than the alternatives. 'Risk-assessment' might be "concessional" – but there is nothing objectionable, in itself, by the inaction it might recommend. So if inaction is to count as complacency, we think it should not be resisted. To reject this sort of 'complacency' is, in effect, to ignore real downside risks in the name of a high minded ideal.

### B. *Is Real-World theory 'anti-aspirational' or 'complacency-inducing'?*

One might nonetheless think that, in some way, real-world theory suppresses aspiration and induces complacency. Perhaps, at least in some people, at some times. Yet, on the other side, ideal theory might well encourage hopelessness and induce disengagement. Which effects one or the other might actually have is an empirical question, about which we have at best anecdotal evidence. We do, though, think it worth pointing out that real-world theory's recognition that not all is possible is fully compatible with serious aspiration – not for the impossible, but for what is possible and worth working hard to do. The large gap between, on the one hand, how things could and should be, and, on the other hand, how things are, is, sadly, room enough for lifetimes of aspiration.

Estlund's 'anti-complacency' speculation may nonetheless strike some readers as a plausible one. Perhaps ideal theory is inspirational. And perhaps real-world theory, by offering a 'reality check' on what might be achieved, serves to emphasize how difficult it may be to secure more justice and, for that reason, dampens enthusiasm. But then two questions arise. One is whether such dampening is always a bad thing. There are at least some contexts in which fervor in favor of an unworkable ideal could use some damping. The second is whether, when inspiration is sought, it might be better supplied by something other than visions of principles of justice – perhaps by specific Utopian *proposals* that might be engaged in with others, rather than by a commitment to a vision of ideal *principles*. Estlund makes it clear that he rejects Utopian proposals on the grounds that they require what is impossible: he is, he says, instead in search of 'principles', which he believes ideal theory can provide. Yet to the extent that the case for the



ideal approach depends on supplying inspiration, substantive Utopian schemes might do a better job than abstract principles.

C. *Psychology and Theory Choice:*

In what ways might any psychological/motivational effects of particular ways of social theorizing be relevant to theory choice? There are two possible targets of concern here: theorists themselves, on the one hand; and the individuals whose understanding of the social order is influenced by the theory in question, on the other. And in each case, the concern might be with general effects of independent normative relevance; or with effects that bear on the theorizing. It is, for example a familiar claim that the study of economics tends to make economists more selfish. That selfishness might be regrettable in itself, and something that they should protect against, without it having any implications for the quality or accuracy of their theories. Alternatively, it may lead to distortions in the attribution of motivations to the subjects of economists' models – leading them to assume that the individuals who inhabit their models are more selfish than they actually are. Or the effects of the economists' claims about the social order might actually change the motivations of the subjects of their models: if, as a result of the economists' models, people come to believe that others are selfish, they may be less inclined to selfless behavior themselves.<sup>9</sup>

We take it that Estlund's concerns are not directed at the induced moral qualities of political philosophers. It is not clear to us whether he is concerned mainly with the impact of theorist's attitudes on the quality of the theories or with the impact of the theories on the attitudes (and actions) of those who accept the theories. To the extent that it is the latter, a couple of observations seem in order. First, the concern seems to depend on political philosophy having a greater influence on popular opinion than is perhaps entirely plausible. Second, in more *ad hominem* spirit, the argument itself seems to be a classic variant of the real-world argument. It takes what it supposes to be an empirical constraint about psychological propensities (to be uninspired by less than ideal principles and become complacent) and suggests an 'optimal response' to that constraint – namely, thinking about justice as it would be if people were perfect. This is a style of reasoning that would be entirely congenial to the real-world theorist – but it looks suspiciously like a self-defeating line for a 'non-concessive' theorist.

Nevertheless, we suspect that this anxiety about complacency is a more significant element in Estlund's project than he allows. Our sense is that he thinks there are important connections between the way one conceptualizes normative concerns and their motivating power. We share that conviction. As we read him, he also thinks that people are *in general* prone to complacency and a general moral torpor – and that securing maximal justice requires that they be inspired to discover the best of themselves, and that the way justice is theorized plays a role in that inspiration project. We are much less confident in this, though we do not deny this possibility. We do, in any case, think we need more than arm chair psychology to understand the ways, if any, in which different ways of thinking about moral principles might influence their motivational impact. Here we simply note that the structure of argument is itself concessive.

D. *Estlund's Example:*

Consider finally an illustrative anecdote that Estlund deploys in chapter 1.

---

<sup>9</sup> There is in play here a distinction relevant for the theorists in question between 'doing theory' and 'promulgating theory'.

*Suppose we are hiking, and we spy a beautiful spot some miles off, down the slope, across the valley. It isn't just beautiful, it looks like a great place to stay, or even to live. Alas, it is not yet clear whether we can get there, so we might try to contain our excitement. Be realistic. Things are fine where we are, so we could just conclude that the new spot is not really worth considering.*

Clearly, we can distinguish three relevant questions here. How much more beautiful is the spot across the valley? How likely is it that, if we tried to get there, we would be successful? What are the costs of trying? It isn't clear on the face of things how real-world theorist and ideal theorists would differ in approaching these questions. But what is clear is that the questions themselves are independent<sup>10</sup> – and that their independence ought to be maintained. Estlund, in discussing this example, seems to assume that, for real-world theorists, the answer to the second question might leak into the first. He seems to be supposing that real-world theorists would recommend our deliberately distorting our judgment of the beauty of the distant prospect because we think we might not be able to get there – or because, turning to the third question, getting there would prove to be too much trouble. No real-world theorist will recognize herself in that picture.

Perhaps, though, people's evaluations of something are influenced by their judgments of whether that thing is hard or impossible to get. The direction of any resultant bias is unclear. People might think better of something because they don't or can't have it, or they might think worse of something because it is hard or impossible to get. Any difference between ideal and real-world theorists does not lie in the direction of bias but in their different approaches to the underlying psychological facts (assuming they are facts). We take it that the idealist would be inclined to insist that any such psychological flaws be ignored: they are after all departures from ideal principles of reasoning. Whereas, we take it, the real-world theorist would give thought to how the errors in question might either be overcome or offset in some way so as to secure better actions.

There is further question lurking: how bad would it be if our attempt to get to the more beautiful location fails? Suppose the route involves walking along a cliff-face path, where if we falter, we plunge to our deaths. Ideally, we should not slip. It would be bad if we did, and if we are careful, we won't. But we might. It seems clear that slipping is a contingency that we ought to reckon with – and the fact that we won't slip if we take adequate care does not in itself ensure that we *will* take adequate care. Furthermore, the ex post judgment that we *should* have taken more care, if we did happen to slip, doesn't seem to meet the case: given what we know of ourselves and our carefulness, trying to get to the more beautiful spot may be a venture that we should not undertake. So much seems unexceptionable.<sup>11</sup>

We expect that Estlund would not disagree. But then, what is the work that ideal theory is supposed to be doing? We turn to this in the next section, when we consider what we take to be his third, and we think the most intriguing, argument.

---

<sup>10</sup> In particular, it is worth emphasizing that the probability of getting there, if one tries, is not the same as the expected costs of making the attempt.

<sup>11</sup> Interestingly, to ignore the prospect of our own carelessness seems to be an instance of...well, 'complacency'.

We see another reason one might worry about real-world theory breeding complacency that Estlund does not mention but that we think worth putting on the table. The worry is prompted by the example, which Estlund does discuss, of Professor Procrastinate:<sup>12</sup>

Professor Procrastinate receives an invitation to review a book. He is the best person to do the review, has the time, and so on. The best thing that can happen is that he says yes, and then writes the review when the book arrives. However, suppose it is further the case that were Procrastinate to say yes, he would not in fact get around to writing the review. Not because of incapacity or outside interference or anything like that, but because he would keep on putting the task off. (This has been known to happen.) Thus, although the best that can happen is for Procrastinate to say yes and then write, and he can do exactly this, what would in fact happen were he to say yes is that he would not write the review. Moreover, we may suppose, this latter is the worst that can happen. It would lead to the book not being reviewed at all, or at least to a review being seriously delayed.

It is easy to think both that (i) the Professor ought to just buckle down, accept the invitation, not procrastinate, and write the review and (ii) knowing his proclivities, the Professor ought to decline the invitation and save everyone from grief he would cause were he to do otherwise.<sup>13</sup> It is also easy to think of (ii) as a sad capitulation to weakness that reflects, if not complacency, at least a failure to challenge himself. More generally, one might think that letting our collective anticipated failures to, say, comply with the principles that shape our understanding of what those principles demand is a similar sad capitulation to weakness.

However, here we think the change from the individual to the group makes a significant difference. Even if we stick with (i) about the professor and think each of us should not capitulate to our own weaknesses, anticipating *others'* weaknesses and adjusting our understanding of what we each ought to do in light of the independent shortcomings of others, strikes us a very different case. There is of course room to press the analogy between the individual and the group. And in some contexts, we are fine with that. But to the extent the standards at issue are applicable to groups of individuals conceived as independent, albeit interacting, agents – as we think principles of justice and criteria for political institutions are – we think the criteria and institutions that apply to all must treat the possible, and indeed likely, shortcomings of others as parameters that help determine what we, individually and collectively, ought to do.

That is not to say that we should take as given and fixed the likely failures of people, treating justice as silent about those failures and leaving us each under no obligation of justice to work to improve our own and other peoples' performance. It is, though, to say that, in contexts of political philosophy, the duties of us all are rightly understood as sensitive not just to what is possible for each individual to do but also to what each can reasonably expect others to do.<sup>14</sup> This allows a theory that might begin by identifying what each ought to do if only others would

---

<sup>12</sup> p. 29 and following and Chapter 8: The Concessive Requirement, pp. 152-153. See Jackson, F., & Pargetter, R. (1986)

<sup>13</sup> There is plenty of interesting debate about this revolving around whether what we ought to do depends on what we will actually do or what we could possibly do.

<sup>14</sup> Here we are thinking that there is an important difference between what we reasonably anticipate someone will do and what we reasonably expect that person to do. But that latter is, on our view, rightly sensitive both to what is easy or difficult for that person to do (which itself is often shaped by what others will do).

do as they ought, but it requires *not stopping there* whenever it is reasonable to expect that others will not comply with the resulting principles. What we actually ought to do in our actual circumstances depends on what those circumstances are actually like, and a relevant feature of those circumstances, in political contexts, involves not just what others could possibly do but also what they are most likely to do.<sup>15</sup>

### V. The Need for a Systematic Theory

We have just said, in effect, that our view might permit *beginning* with ideal theory. The third argument we want to consider maintains that we *must* rely on ideal theory, if we are to develop a systematic and principled theory of justice and of political institutions more generally.

This third argument is more cited in passing by Estlund than articulated or defended. It claims that no systematic or principled theory of justice is available in the absence of ideal theory. As Rawls puts it

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 [posed by an imperfect world]...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.<sup>16</sup>

Exactly why we should believe ideal theory is needed for this purpose is unclear.<sup>17</sup> We will explore some possibilities. What *is* clear is that Rawls thinks sorting through different real-world theories concerning what justice demands in circumstances shaped by our shortcomings ultimately requires appeal to the principles we would discover only in doing ideal theory. This, if true, would be a powerful argument for ideal theory and against those who think that all the important theorizing about justice can, and probably should, be ‘real-world’.

Clearly, with respect to at least some evaluative/normative concepts, a substantial and informed, if not principled and systematic, grasp of them does not require any appeal to the ideal. Consider Estlund’s parable of a beautiful destination in the distance from which we briefly quoted earlier, and consider specifically its beauty. Rawls’ (and, we are thinking by implication, Estlund’s) recipe for determining what beauty *is* and/or what it might require or call for, is that we start by trying to conceive of the *perfectly* beautiful, and then compare locations by reference to how far each falls short of that ideal. But that recipe seems unpromising, to say the least. We can surely make aesthetic judgments of a purely comparative kind (for example, that the distant location is more beautiful than the one we occupy -- and perhaps very considerably so) without having any notion of what ‘ideal beauty’ would consist in and even without any conviction that such a notion makes any sense. The thought that we cannot know anything about what beauty really *is* or how to advance it unless we can specify what the ideal is, seems plainly mistaken. And what

---

<sup>15</sup> To underline the distinction between the reflexive and the many person case, consider what the editor should do. Should she send the paper to Procrastinate anyway, knowing what she knows about his most likely response? Should she instead send the reviewing exercise to the next best person on her list – Professor Prompt? If she chose the latter option, she might be letting Procrastinate off the hook – but is putting him *on* the hook by making his failures more conspicuous, a consideration that the editor ought to weigh in securing the predicted best outcome? Whatever the answer to these questions, there is no implication that the editor should feel ‘complacent’ about Procrastinate’s weaknesses – she might well feel frustrated by, and disgusted at, his professional slackness!

<sup>16</sup> Cited by Estlund on p. 7; from Rawls’ *A Theory of Justice*, revised edition, section 2, 8.

<sup>17</sup> Estlund mentions Spencer as having anticipated Rawls, and Sidgwick as having resisted, while registering that they were both concerned with what sort of theory would be most interested in whether ideal theory, or real-world theory, would actually be most useful.

is true of beauty is also true of other normatively relevant notions, such as well-being: judgments of what might make us more and less well-off can be undertaken without any conception of what an ideal level of well-being would be. Indeed, there is a tradition in political theory that begins with a very bad state of affairs and then attempts to derive a metric of normative ‘betterness’ based on departures from that bad state. This is a tradition we associate with Thomas Hobbes. Hobbes’ procedure is, as we see it, the mirror image of Estlund’s, itself starting with an extreme we might actually never have faced (given the natural ties of family and attraction) and in any case do not now face, but applying at the negative rather than the positive end of a broad spectrum. Just as Rawls supposes that the true nature of justice and political association cannot be understood without appeal to what the best might conceivably be, Hobbes thinks it cannot be understood without appeal to what the worst might conceivably be. Both start where we are not, with the hope that doing so will help us understand where we are. We do not necessarily doubt the value of either approach. But we think neither is necessary for understanding what would be boons, or disasters, or what are improvements, in the relevant domain.

To reiterate the claims in section III, one can have a measure – a principled and systematic measure -- of a normative concept without that measure having a limiting value (in either direction).<sup>18</sup> Any such measure of justice, if one is available, will tell us whether one situation is more just than another. That information may not be enough to guide choices when there are risks associated with any course of action – such choices require a measure of how much worse one outcome is than another. But that information is not provided by the knowledge that both fall short of the ideal – if there *is* an ideal! Moreover, if someone is a reliable judge of which across a broad range of situations is the most just, they seem to have a reasonable claim to understanding justice. Contrast that person with one who can specify an ideal of justice but cannot tell in any situation whether one or another actual situation is more or less just.

So far, of course, one might grant these points and yet hold that the understanding on offer is, in important respects, not systematic, even if, say, useful. What might make that plausible? Perhaps, justice is a concept like equality (or telling the truth or the straightness of lines) that is defined by its limiting realization -- and precisely not like beauty or well-being or pleasure, which may lack such a limiting realization. Maybe. But to settle that question would seem to require us to say a good deal about why that would be true of justice than Rawls or Estlund have said.<sup>19</sup>

Estlund identifies two interesting grounds for thinking that ideal theory is the place to start. One is that “concessive requirements arise only from failure to meet the non-concessive” (p. 31). The other is that on the fully concessive view “there is no single salient standard of social justice at all, just a field of concessive requirements” (p. 31) We have doubts about the force of each.

With Socrates, Hobbes, Hume, and Rawls, we see justice as having its place and point thanks to what Rawls refers to as our facing the “circumstances of justice” – which Hume helpfully characterized as depending on limited benevolence, limited scarcity, and the opportunity for mutual advantage through cooperation. As against Rawls, though, we see considerations of what is feasible – specifically, considerations concerning what people, generally, might *willingly* do -- as often a crucial factor in understanding the relevant circumstances and so in identifying what justice would require people, individually, to do. What you or I are required to do, by justice,

---

<sup>18</sup> This claim we associate with Amartya Sen (2009) – and John Broome (1999).

<sup>19</sup> Moreover, as earlier emphasized, there are many measures of inequality or lack of straightness available – and specifying the ‘correct’ one is necessary for a complete (as well as an applicable) understanding of equality/straightness in any world where equality/straightness is problematic.

depends, on this view, on what *others* would willingly do.<sup>20</sup> Wishing away peoples' unwillingness to do certain things, in the name of a frictionless theory of justice, strikes us as losing sight of a key factor in understanding how people should see their own duties in light of what others will, or will not, do. If this is right, the requirements that Estlund views as concessive do not "arise only from failure to meet the non-concessive." Rather, they emerge directly from an appreciation of the circumstances that give justice its point.

This seems especially true in terms of what *institutions* might be put in place to secure justice. Suppose that if 100% of people conform, the best institutional package is A. But if 80% of people comply, the best institutional package is B. If 80% compliance is what can reasonably be expected, then (we think) that B is what justice requires (even as it predictably falls short of what a concern for justice would lead us to wish for). In the case of individual behavior<sup>21</sup>, agents determine their level of compliance (whether 100% or zero or perhaps something in between, to the extent compliance is under their control): but in the case relevant for 'political action' single agents do not themselves determine the level of compliance – each can only try to influence what others will do. The point here is that what *all* should do, assuming all do as they should, is not in general the same as what *each* should do except in the special case of full compliance.<sup>22</sup> This distinction leads us to suspect that when it comes to thinking through what justice requires of each of us we should not typically be thinking about what it might, in ideal circumstances, require of us all.

It also means that the non-concessive theory's claim to non-arbitrariness looks suspect. There may be only one set of requirements for ideal conditions, rather than the many sets that emerge depending on which less than ideal circumstances we consider. Yet, to the extent the question is which requirements apply to us, assuming away the conditions we actually face is no less arbitrary than taking them into account. Concerns about arbitrariness do not disappear, but they do face any view that privileges some conditions over others (as ideal and real-world theories alike do).

Our working hypothesis is that a fully justified, and systematic, theory of justice needs to take account not simply of what is physically possible, but what is possible through and because of what people are willing to do. The idea is that what one person or group is required by justice to do is sensitive to what others can reasonably be expected to do willingly (albeit, sometimes, in the light of rewards or punishments imposed by people who *can* reasonably be expected to administer those in the interests of greater 'justice').<sup>23</sup> The overarching view is that the principles of justice are meant to govern the potentially mutually beneficial interaction of individuals and groups. To put to one side the constraints on what might actually be possible

---

<sup>20</sup> In tilting this way, we are embracing what Estlund thinks of as "the Bent View" of human nature. But we draw a distinction, that Estlund does not, between seeing the principles of justice as shaped, in some way, by what people might willingly do and seeing them as shaped, in some way, by what motives are *characteristic* of human agents (p. 103). Just what counts as what people, or a particular person, might willingly do we here leave unexplored, even as we think it is neither set by nor restricted to characteristic motives of human agents. See Estlund's Chapter 7 "Justice Unbent," where he helpfully distinguishes, as we think must be done, between individual motives being individually mitigating and their being collectively mitigating (p. 124 and following).

<sup>21</sup> Of the kind familiar from the antics of Professor Procrastinate and his commentators.

<sup>22</sup> Of course, you do not unilaterally determine what institutional arrangement will prevail. And since your vote (or promotional activity) makes very little difference to the outcome, you might be inclined to vote in an exemplary fashion for institution package A. (See Brennan and Lomasky (1993) and Brennan and Sayre-McCord (2015) for relevant discussion. Such complications are however not essential here.)

<sup>23</sup> That is, we do not suppose that with heterogeneous actors selection cannot do important normative work. Some institutions work by incentives; some by selection; most by some degree of both. In other words, what 'people' are willing to do does not presuppose that all people are identically motivated — *which* 'people' is often a central question.

through the agency of individuals and groups is, we think, to ignore the underlying insight of Socrates, Hobbes, Hume and Rawls that justice is best understood as a solution to a shared problem we would otherwise face to living together in ways that are mutually beneficial. What others might do through their agency is a key opportunity and a key constraint on what it might be possible for people to accomplish.

More, of course, might be possible than what people might (individually or collectively) accomplish, but what lies beyond that border is, we think, beyond what falls within the scope of the demands of justice. If this is right, the unwillingness of others to comply with what they take to be the demands of justice is relevant to what the demands actually are.

What are those limits, and how might we determine them? They certainly go beyond what is physically possible, as we have noted. But how far, and in which directions, is not at all clear. People are sometimes capable, under extreme circumstances, of more than anyone imagines possible. At the same time, others manifestly fail to be able to get themselves to do even what they deeply desire to do. The limits are, we think, real, but are neither easily discoverable nor the same for everyone.

In thinking about the demands of justice, therefore, we hold that the reasonable approach is to work with what we can reasonably expect the limits to be, in general. Doing so, though, predictably generates a different view of what principles of justice might require than would imagining ideal conditions in which all willingly comply with the principles on offer.<sup>24</sup>

## VI Summary and Conclusions

Our approach to Estlund's *Utopophobia* has been to take what we see as three of the most significant arguments he offers against Utopophobia and react to them.

The first is, as we read it, a combination of "let a thousand flowers bloom" and "it ain't necessarily so" – both amounting to an appeal to the onus of proof. Estlund thinks that real-world critiques wipe ideal theory off the table without establishing a decisive case against it. In that sense, he thinks that the real-world (non-ideal) critique contains an element of 'phobia' (a distaste that is ultimately irrational). We are not unsympathetic to this line. But we are inclined to think that, in the absence of a decisive case for a proposition or its opposite (p and ~p), it is acceptable to go with the balance of argument. In that setting, to describe as a "phobia", a reasoned judgment that real-world theory has more going for it than the ideal theory alternative, strikes us as rhetorical overreach.

Estlund's second argument is that real-world (or 'concessionist') theory encourages complacency – and that it lacks the aspirational character that he thinks all normative judgments do or should contain. In fact, we suspect that this argument is an important element in understanding Estlund's position (as important psychologically as it is argumentatively) – and we have therefore given it reasonably detailed attention. Ultimately, we are unpersuaded. We doubt that real-world theorizing leads to complacency (however precisely that term is understood); and though complacency is a danger, so, as we see it, is the moral recklessness ideal theory might correspondingly encourage. Moreover, but mostly as a side point, we think

---

<sup>24</sup> Estlund helpfully introduces and explores the idea of (what he terms) Plural Requirements, which do not apply to an agent (considered alone) but only to sets of agents and then impose obligations on each only on the condition that the other(s) act in certain ways. See Chapter 12. A nice question that we here leave to one side is whether these Plural Requirements are themselves sensitive to what people might willingly do.

this a surprising argument for Estlund to embrace, since it seems in its form to be a variant of the real-world/concessive approach about which he declares his scepticism.

The final argument is perhaps less Estlund's than it is John Rawls's -- and it is not entirely clear to what extent Estlund endorses it. Yet it is one that we would find compelling, if it were right. It is the argument that a systematic theory of justice requires an appeal to the principles identified in doing ideal theory. We think this claim mistaken. For one thing, there are some normative concepts that clearly do not admit a limiting case – and the claim that one cannot therefore systematically understand what they amount to seems misplaced. So, idealization cannot be necessary for a systematic understanding of all normative concepts. Yet this is of course compatible with thinking that justice does admit of a limit and that a systematic understanding of justice requires an appeal to the principles appropriate to that limit. We acknowledge this possibility but we doubt that it is true. Impressed by the idea that the concept of justice is best understood in terms of the shared problems we face together, we suspect that the concept of justice requires, from the start, an appreciation of the problems we face when circumstances are less than ideal.

In any case, even if justice does admit of a limit, we think it is clear that an account that appealed only to the principles appropriate to the limiting case would not constitute an appropriately systematic theory of justice. Indeed, on Rawls' view, no less than ours, a full understanding of a justice (even assuming it does have a limiting case) requires working out in situations where it is not fully realized, which is the more/most just: precisely why and in what respects one situation is less just than another seems to us to be an important element in the understanding of the *concept*, as well as a necessary feature for its application in the imperfect world we face.

### Bibliography

Brennan, G. and L. Lomasky (1993) *Democracy and Decision* Cambridge University Press: New York

Brennan, G. and G. Sayre-McCord (2015) "Voting and Causal Responsibility" *Oxford Studies in Political Philosophy* 1: 36-59

Broome, J. (1999) "Goodness is Reducible to Betterness" in *Ethics out of Economics* Cambridge University Press: Cambridge ch 10

Cohen, G. (2003) "Facts and Principles" *Philosophy and Public Affairs* 31: 211-245

Jackson, F. and R. Pargetter (1986) "Oughts, Options and Actualism" *Philosophical Review* 95: 233-255

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

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

Taylor, S.E.; Brown, J. (1988) "Illusion and well-being: A social psychological perspective on mental health". *Psychological Bulletin*. 103 (2): 193–210.



