

Gopal Sreenivasan

## ‘Oh, but you should have’: Estlund on Normative Consent

*Democratic Authority* is a tour de force. Among its considerable virtues, David Estlund’s book manages to introduce genuinely new and plainly important ideas on well-worn topics right at the heart of modern political philosophy. Indeed, it does so not once, but several times. In this discussion, I shall concentrate on just one of those ideas, the idea of ‘normative consent’.

Estlund offers his idea as a novel justification of political authority—ultimately, as a justification for the authority of democracies. But, officially, the justification is structured as an argument by analogy, with the analogy running from the authority of a jury system to the authority of a democracy. Accordingly, I shall focus here on Estlund’s point of departure: the normative consent justification for the authority of a *jury system*. Not only does this narrower focus contain all the essential points that need to be considered, but it also absorbs the bulk of Estlund’s own attention in chapter 8.

1. The central claim behind the notion of normative consent is that *refusals* of consent are sometimes disqualified or null. Estlund opens the door for this claim by observing an interesting asymmetry within standard consent theory. As he represents it, the standard theory holds that

without consent there is no authority (the *libertarian clause*), but unless there are certain nullifying conditions (the *nullity proviso*), consent to authority establishes authority (the *authority clause*) (119).<sup>1</sup>

The ‘nullity proviso’ derives its content from some distinction between qualified and disqualified consent. Yet whereas the power of consent to establish authority is limited by this proviso (disqualified consent fails to

<sup>1</sup> Parenthetical page citations refer to D. Estlund, *Democratic Authority* (Princeton: Princeton University Press, 2008).

establish authority), the power of non-consent to exclude authority is *not* limited by any such proviso.

Estlund's distinctive suggestion is that *non*-consent can sometimes be disqualified, too. When it is disqualified, "the authority condition is as it would have been if there had been qualified consent" (123). That is to say, authority then obtains. Hence, contrary to voluntarism, "authority can simply befall us, whether we have consented to it or not" (117). Nevertheless, as Estlund is anxious to emphasise, authority "retains *some* connection to the will" (130) on the resultant theory of normative consent. This contrasts with theories of 'direct authority', which sever all connection between authority and the individual's will. Estlund restricts himself to a modest version of normative consent theory (132), thereby leaving it open whether there are any direct bases of authority (in addition to the bases of consent and normative consent).

Somewhat unusually, Estlund defines 'political authority' in such a way that there can be (rightful) political authority without political legitimacy (134).<sup>2</sup> (This makes his definitions structurally like Hobbes's and unlike Locke's or, I should have thought, the standard definitions.) *Authority* is defined as "the moral power to require action" (119), while *legitimacy* is defined as "the permissibility of coercively enforcing commands" (134). They are further distinguished by the fact that claims of legitimacy, but not those of authority (134; but cf. 120), are burdened by the 'qualified acceptability requirement' that Estlund introduces in chapter 3. Briefly, this means that justifications of legitimacy can succeed only if they do not rely on doctrines that 'qualified views' could reject.

We are now equipped to review the basic application of Estlund's theory of normative consent. It proceeds in two steps. To begin with, Estlund argues that an effective jury system *has* political authority (within its jurisdiction). He then goes on to argue that the moral basis of this authority can be supplied by an appeal to normative consent. Let us consider these steps in turn.

The specific instance of political authority with which Estlund credits an effective jury system is the moral power to require of individuals at large that they exonerate defendants (or not) as a duly constituted jury has decided (140–41). For example, individuals acquire a duty not to punish

<sup>2</sup> It seems to me that this is a serious mistake, but I shall not pursue the matter. It is related (tangentially, rather than crucially) to a few points raised below.

any defendant privately whom such a jury has declared innocent. Estlund maintains that an effective jury system's possession of this power follows from the fact that its publicly visible adequacy satisfies the antecedent of the 'antivigilante principle'. This principle states that "when there is a system that serves the purposes of judgment and punishment without private punishment, then private punishment is morally wrong" (140).

Estlund's account of the moral basis of this authority is complex. It weaves together several strands of argument and it may help to think of this weaving as taking place under a pair of nested umbrellas. The inner umbrella covers Estlund's districting solution to a humanitarian problem, while the theory of normative consent functions as an outer umbrella, covering the inner umbrella.

At the centre of his analysis is the idea, familiar from traditional accounts of the state of nature, that 'juridical anarchy' is a humanitarian problem of great urgency, to which the establishment of an effective jury system represents a partial solution (146–47). Estlund holds that the *urgency* of this problem imposes humanitarian duties on everyone to contribute to its solution. However, he also regards the required individual contributions as being *limited to a fair share*. An individual's fair contribution to eliminating juridical anarchy includes refraining from private punishment of defendants declared innocent by a duly constituted jury. Estlund argues, further, that while juridical anarchy is a *global* problem, the duty to refrain from private punishment can be suitably tied to the verdicts of *local* juries by treating the division of the world into districts (i.e., jurisdictions) as part of the global solution. Finally, drawing on his innovative discussion of *epistemic proceduralism* in chapter 3, Estlund insists that the urgency of a humanitarian task does not justify any old solution. An important part of the distinctive merit of the jury system solution rests on the combined facts that it has "some decent tendency" (137) to produce correct verdicts and that it is immune to objections from qualified points of view (139).<sup>3</sup> At its strongest, the claim adds that no qualifiedly acceptable arrangement has a better tendency to produce correct verdicts.

Yet, for all that, it remains one thing for individuals to be duty-bound to refrain from private punishment and another for them to be (rightfully)

<sup>3</sup> In the end, the analogy Estlund wants to defend between democracy and a jury system is meant to hold across all of the strands highlighted in this paragraph (157).

subject to *authority*, specifically. So the inner umbrella still leaves something out. To complete the analysis, Estlund argues that—given an effective jury system in their jurisdiction—individuals are not only duty-bound, but would also be *wrong not to accept a duty* to refrain from private punishment. Moreover, this supplementary conclusion can be given the same moral basis as the conclusion that individuals are duty-bound to refrain from private punishment.<sup>4</sup> Now recall that, on the theory of normative consent, the nullity of non-consent inaugurates the same authority condition as qualified consent does. Since an individual's qualified consent to a duty to refrain from private punishment would vest the jury system with authority (to forbid private punishment), so too does the nullity of her non-consent to that duty. Ultimately, then, Estlund locates the moral basis of an effective jury system's authority in the fact that non-consent to a duty to refrain from private punishment would simply be null.

2. Before turning to my principal criticism, it is worth pausing to notice two significant lacunae in the overall argument to which Estlund himself draws attention. One concerns the definition of authority, while the other concerns Estlund's distinction between qualified and disqualified non-consent. Let me start with the former.

Recall that Estlund defines 'authority' as the moral *power to require action*. When pressed, this formulation turns out to be ambiguous as between the 'power to command action (rightfully)' and the 'power to do things that result in others' being subject to moral requirements on action'. Only the first power constitutes genuine authority. Following Estlund, we can call

<sup>4</sup> This is fairly straightforward provided that accepting a duty to refrain from private punishment can itself be made out as a suitable contribution to the task of eliminating juridical anarchy (in one's district). While I see no particular obstacle to making this case out, Estlund himself does not pursue this option. Instead, he emphasises (and is perhaps overly impressed by) the undeniable fact that the contribution to eliminating juridical anarchy made by *<accepting a duty to refrain from private punishment>* is distinct from that made by *<actually refraining from private punishment>*. This leads him to introduce a different urgent task to absorb the contribution made by *acceptance of the duty* (the 'general commitment task' [152–53]), and thereby to distinguish somewhat the moral bases he supplies for the two conclusions discussed in the text. It seems to me that this move is neither necessary nor altogether satisfactory. But it does mean that my statement in the text is not perfectly accurate as a description of Estlund's own position.

cases in which the second power is exercised *side-effect* cases (143). He offers the example of a brutal dictator with a petulant child: If the dictator's ministers do not do as the child tells them, the dictator will get angry and "unleash brutality on the masses" (118). The ministers are therefore morally required to do what the child tells them, but the child nevertheless lacks authority.

Hence, from the fact that a jury system has the power to require action, it does not follow that it has authority, as the possibility always remains that the jury system's exercises of that power are mere side-effect cases. Estlund freely admits that he does not know how to license this inference (143). As a result, he can only *assert* that a jury system's exercise of its power to require action is a case of political authority (144). For what it is worth, I doubt that an effective jury system has political authority,<sup>5</sup> but I do not think that this point is actually very damaging. After all, it is still perfectly clear that *democracies* have political authority. So if jury systems lack it, all that is really affected is the surface organisation of Estlund's argument (for the authority of democracy) as an argument by analogy. Its fundamental logic remains intact.

The second lacuna is more serious. By definition, the *nullity* of non-consent to some state of affairs entails the authority condition that would result from qualified consent to that same state of affairs. However, that tells us nothing about *when* non-consent is null. All it tells us is something about what follows, given that non-consent is null. Unfortunately, the immediate conclusion of the outer umbrella argument is *not* that non-consent to a duty to refrain from private punishment would be 'null', but rather that individuals would be *wrong* not to consent to that duty. Worse, Estlund explicitly denies that the nullity of (any particular) non-consent follows from its wrongness. Indeed, where sexual offers are concerned, for example, he insists that non-consent can be wrong *and yet not* null: wrongful non-consent to sex "retains its prohibitive force" (126).

<sup>5</sup> My doubts have to do with the highly constrained character of the 'authority' with which Estlund credits a jury system. It strikes me that *if* a jury system has political authority, it is only delegated authority rather than original authority. As Estlund describes them, jury systems have (as it were) the power to spring traps that are there anyhow, but not to lay any traps of their own. Less colloquially, but more abstractly: jury systems can satisfy the antecedents of pre-existing conditionals, but cannot impose any novel consequents of their own.

Here, too, Estlund simply concedes that he does not know how to specify when wrongful non-consent is null (127). He conjectures that it is null whenever the nullity of wrongful non-consent would not permit interference with the non-consenter's person or property. On this basis, he invokes the terms of his distinction between authority and legitimacy to argue that qualified consent to *authority* does not actually "permit anyone to do anything" at all (127)!<sup>6</sup> A fortiori, it does not permit interference with person or property. Neither, then, would the nullity of wrongful non-consent to authority.

3. Each of the self-diagnosed lacunae involves an unlicensed inference. I should now like to argue that Estlund's normative consent justification for authority is beset by a further, and more damaging, non sequitur. To simplify matters, let us assume that there is some tenable ground on which to identify the sub-set of cases in which wrongful non-consent *is* null. For the remainder of the discussion, our attention will be restricted to this sub-set. (We may therefore ignore the second lacuna.)

Consider the following conditions, where X and Y are individuals and  $\phi$ -ing is an action:

- (a) X  $\phi$ 's Y.
- (b) Y does not consent to X's  $\phi$ -ing, but Y *should have*. (Y's non-consent is wrong).
- (c) All things considered, it is nevertheless permissible for X to  $\phi$  Y.

As background, let us assume throughout that, ordinarily, (a) would be wrong. But (a) would also become permissible if Y consented to it. What justifies (c)?

One justification for (c) is the normative consent justification. It treats the fact that Y *should have* consented to X's  $\phi$ -ing as entailing the normative upshot that would have resulted had Y actually consented (i.e., it treats Y's non-consent as null). Since Y's actually consenting to X's  $\phi$ -ing would have made X's  $\phi$ -ing Y permissible, (c) is justified. Schematically, (b) justifies (c).<sup>7</sup>

<sup>6</sup> To my mind, this amounts to exploiting one difficulty in the argument to escape another.

<sup>7</sup> Strictly speaking, this is a generalisation of Estlund's normative consent justification. It employs the broader notion of a 'normative upshot', where he uses the notion

However, another possibility is that the permissibility of (a) is justified, fundamentally, by the fact that it is much better, all things considered (ATC), if X  $\phi$ 's Y than if X does not. Moreover, this same underlying fact justifies not only the permissibility of X's  $\phi$ -ing Y, but also the fact that Y should have consented to it. Schematically, a more fundamental fact justifies both (b) and (c). Call this the *riding roughshod* justification, as it involves the idea that the ATC ought can ride roughshod over the significance of Y's consent.<sup>8</sup>

It may help to compare these justifications in a garden variety illustration of conditions (a)–(c), as a prelude to revisiting Estlund's nested umbrella argument. Consider a case of emergency trespass,<sup>9</sup> where Y is a landowner and X is the father of an injured child:

- (a\*) X crosses Y's land to bring his injured child to hospital.
- (b\*) Y does not consent to X's trespass, but Y should have.
- (c\*) ATC, X's trespass is nevertheless permissible.

Say the injured child will survive if and only if the father commits the trespass (there is no other timely route to a hospital). Then it is (very) plausible to see the (value of the) child's life as an underlying, more fundamental fact that justifies both (c\*) the ATC permissibility of the father's trespass and (b\*) the wrongfulness of the landowner's non-consent.

Let us stipulate that (a\*)–(c\*) are all truths. The normative consent justification and the riding roughshod justification represent competing explanations of the truth of (c\*). A crucial difference between them is that the normative consent justification of (c\*) assigns a contributing justificatory role to (b\*), the wrongfulness of the landowner's non-consent, whereas the

---

of a '[resultant] authority condition'. A resultant authority condition is a species of normative upshot.

<sup>8</sup> Notice that the riding roughshod justification need not hold that Y's consent has *no* significance. It can require what is (relatively) at stake to cross some threshold, for example, before the threat of riding roughshod materialises. Below that threshold, Y's consent is simply determinative. Still, consent clearly has less significance here than it does conventionally. Does it have less significance here than in the normative consent framework?

<sup>9</sup> I borrow this example from J. Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), 98. It involves interference with property, but our simplifying assumption did not commit us to *Estlund's conjecture* as the criterion of when wrongful non-consent is null.

riding roughshod justification does not. On the former, (b\*) belongs to the explanans, whereas on the latter (b\*) belongs to the—strictly, to a separate—explanandum. Since both justifications count (b\*) as true, its truth cannot decide between them.

This brings me to my principal criticism, which comes in two versions. The strong version objects that the riding roughshod justification is actually more plausible than the normative consent justification. At least in the emergency trespass case, I do think that is right. But I need not insist on this. A weaker version of the criticism objects only that the correctness of the normative consent justification *does not follow* either from the truth of (a\*)–(c\*) or even from its 'fitting' those facts—that is, not even from its (apparently) justifying (c\*) and cohering with (b\*). It does not follow because exactly the same claim can be made for the riding roughshod justification; and the two justifications are inconsistent. Hence, to privilege the normative consent justification, we would also need some basis for excluding the riding roughshod justification.

Now Estlund's normative consent justification for the authority of an effective jury system can itself be modelled as an instance of our conditions (a)–(c). Here X is the executive branch in a given district and Y is any inhabitant of the district. Assume that the district already has an effective jury system.

(a\*\*) X prevents Y from punishing privately.

(b\*\*) Y does not relinquish her privilege to punish privately, but Y should have.<sup>10</sup>

(c\*\*) ATC, it is permissible for X to prevent Y from punishing privately.

Estlund's nested umbrella argument first marshals the contents of the *inner* umbrella to justify (b\*\*), the claim that Y should have relinquished her privilege (i.e., consented to a duty not) to punish privately. Whereupon it marshals the contents of the *outer* umbrella to create a role for (b\*\*) in justifying (c\*\*),<sup>11</sup> which it does by claiming that Y's non-consent is not merely wrong, but actually *null*.

<sup>10</sup> In line with traditional accounts of the state of nature, Estlund assumes that individuals *have* a privilege to punish privately (to begin with, anyhow) (142). Note that Y's relinquishing her privilege to punish privately is equivalent to Y's consenting to a duty not to punish privately.

<sup>11</sup> The second lacuna looms here, as an obstacle to opening the outer umbrella.

Alternatively, however, a riding roughshod justification can deploy the contents of the same inner umbrella to justify (c\*\*) directly.<sup>12</sup> It can also deploy them separately, to justify (b\*\*) the claim that Y should have relinquished her privilege to punish privately. (Here the contents of the inner umbrella are analogous to ‘the child’s life’ in the emergency trespass case—both occupy the role of the underlying fundamental fact.) On the riding roughshod justification, the outer umbrella is therefore otiose.

My principal criticism of Estlund’s normative consent justification can be recapitulated briefly. Even if we accept all of its premisses—the truth of (b\*\*) and (c\*\*), the inner umbrella argument as the justification of (b\*\*), and the coherence of normative consent theory—it does not follow that (c\*\*) is in fact justified by normative consent.

4. In closing, let me examine a pair of replies that might be entered in Estlund’s defence.<sup>13</sup> The first reply assimilates my criticism to the ‘direct authority objection’, which Estlund entertains in chapter 7. This objection claims that “whenever it would be wrong to consent to authority in light of certain facts, those same facts already establish authority independently of anything about the duty to consent” (129). Estlund eventually distinguishes two versions of this objection, depending on whether the ‘certain facts’ are required to constitute a ‘sufficient moral basis’ for authority or may instead be any old facts that materially entail a condition of authority (including, e.g., a set of merely physical facts about masses and forces). He calls the former version “more formidable” (130), but it is obscure to me exactly what his answer to it is. It is clear, however, that he is imagining the ‘sufficient moral basis’ proposed by the objection to be altogether different from that proposed by his normative consent justification.

I shall offer two brief rejoinders. First, it should be clear that the ‘fundamental facts’ to which the riding roughshod justification appeals *are* moral facts, and hence do purport to offer a ‘sufficient moral basis’ for authority. Second, far from being altogether different, the sufficient moral basis for (c\*\*) proposed by the riding roughshod justification is a subset

<sup>12</sup> Or perhaps indirectly, by justifying the claim that Y has a duty to refrain from private punishment. The essential point is that the alternative justification does not run through (b\*\*).

<sup>13</sup> Estlund made both replies during the symposium at the Hebrew University.

of Estlund's own! In effect, the alternative I suggest proposes to do all the justificatory work with his inner umbrella and thereby to dispense with his outer umbrella.

The second reply accuses me of misrepresenting the normative consent justification in certain respects. Most significantly, it complains that condition (c\*\*) includes a 'permission to *prevent* Y' from punishing privately. This precisely fits Estlund's definition of *legitimacy*, which concerns the permissibility of coercive enforcement. But his theory only claims to analyse authority.

While this reply certainly has greater merit, I do not believe it affords an adequate line of defence. This time I shall offer three rejoinders. To begin with, the formulation of (c\*\*) could be changed, substituting 'forbid' for 'prevent'. This would move (c\*\*) out of the scope of Estlund's 'legitimacy', and closer to his 'authority', without altering the main point. But there are two problems with this. One is that the parallel substitution of 'forbid' into (a\*\*) yields something that is *not* ordinarily wrong, thereby obviating the need for Y's consent in the first place. Another is that, strictly speaking, the substitution still would not produce a case of 'authority' in Estlund's sense, as he insists that authority does not even convey permission "to issue commands" (127). I believe both problems are symptomatic of grave difficulties with Estlund's definitions (cf. notes 2 and 6).

Next, we could avoid a lot of definitional grief by switching the focus from juries to the predicament of Joe and the flight attendant, who appear briefly in chapter 7. The question is whether a passenger, Joe, comes under the authority of a flight attendant who is trying to direct emergency efforts after a crash. Joe is described as wrongly refusing to consent to her authority. Estlund writes:

you think that Joe has not escaped the authority by refusing to consent. So he is under authority even without having consented. In this case, non-consent to authority is null. (124)

This passage illustrates perfectly the non sequitur I mean to criticise (albeit in miniature, and so without all the rich detail of Estlund's main line of argument). The last sentence does not follow. By appealing simply to the emergency conditions—on the model of the trespass case—a riding roughshod justification can be offered for the first two sentences in the passage, as well as for the wrongfulness of Joe's non-consent. So something

further is needed to privilege the distinctive justification implicit in the claim that Joe's wrongful non-consent is 'null', in Estlund's technical sense.

Finally, and most importantly, my criticism was actually directed at a generalisation of Estlund's normative consent justification. It employed the broader notion of a 'normative upshot' in place of his '[resultant] authority condition'. This is relevant for two reasons. First, changing an action's status from 'wrong' to 'permitted' *is* a perfectly good example of a normative upshot (one equivalent to an exercise of consent, no less), even if the action is a coercive interference. Secondly, the generalisation subsumes Estlund's version as a special case. How can a structural fallacy afflicting the general case not apply to the special instance? At a minimum, it seems reasonable to require an explanation. What is it about 'authority', as Estlund has defined it, that rehabilitates (if, indeed, it does) an inference that we have seen to be invalid when attempted with other normative upshots of consent?

Contrary to what one may have imagined, there are no books of philosophy free from error or deficient argument. So the intuitively appealing contrast between good books and mistaken ones is a chimaera. The salient contrast is rather between the books whose mistakes and defects are *worth* examining and learning from, and the books that are probably best ignored. While I have concentrated on criticism—indeed, because of it—I hope it is abundantly clear that I place Estlund's *Democratic Authority* in the first class.

*Duke University*