

Motivational Limitations on the Demands of Justice

David Wiens

Abstract. Do motivational limitations due to human nature constraint the demands of justice? Among those who say no, David Estlund offers perhaps the most compelling argument. In this paper, I show that Estlund's argument rests on an ambiguous analysis of the concept of ability. Further, I argue that the most plausible specification of his analysis yields the conclusion that at least some motivational limitations—“good faith” motivational limitations—constrain the demands of justice. In fact, my argument against Estlund implies something stronger; namely, that the demands of justice are constrained by what people are sufficiently likely to be motivated to do. Thus, contrary to the prevailing wisdom, it is the business of ideal theory—not just nonideal theory—to work with the motivational capacities people are likely enough to have.

Are the demands of justice subject to feasibility constraints? Put differently, do the “oughts” of justice imply the “can” of feasibility? While numerous philosophers concede that they do (on some interpretation of “ought” and “can”), there is a chorus of theorists who claim that justice remains (largely) untouched by at least one important kind of feasibility constraint—namely, human motivational deficiencies. David Estlund is perhaps the most forceful exponent of this view, arguing that the requirements of justice are *never* blocked by (non-pathological) motivational deficiencies, even if they are hard-wired into human nature.¹ If Estlund is correct, then justice can require us to do things that we *cannot, by nature*, bring ourselves to do. This is a striking result. I shall argue that it is mistaken—motivational deficiencies can constrain the demands of justice.

Section 1 gives some general background and refines the proposition in question thus: a directive is a valid demand of justice upon action (i.e., “ought”) only if some individual or group has the motivational resources required to satisfy the directive (i.e., “can will”). To provide a focal point for my argument, section 2 reviews the important features of Estlund's argument to the contrary. Sections 3 and 4 take Estlund's analysis of

Author's note. Thanks to Dave Estlund, Holly Lawford-Smith, and John Maier for helpful comments. Support from the Australian Research Council (Discovery Grant DP120101507) is gratefully acknowledged.

1. David Estlund, “Human Nature and the Limits (If Any) of Political Philosophy,” *Philosophy & Public Affairs* 39, no. 3 (2011): 207–237, esp. at 228. The “non-pathological” qualification is important, as Estlund grants that “clinical” cases, like addictions, phobias, and compulsions can be requirement-blocking (219). I leave this qualification implicit for the remainder of the paper.

“ability” as a starting point (for reasons to be explained below) and show that motivational deficiencies can undermine abilities and, in turn, block moral requirements under at least one common circumstance — namely, that the motivationally-deficient agent makes a good faith effort to overcome her deficiency. Section 5 shows that my argument in prior sections implies a stronger claim — namely, that the demands of justice depend, to some extent, on how people are *sufficiently likely* to be motivated, not just on how they can be motivated.² It follows that the distribution of good faith motivational deficiencies we find in the actual world constrains not only the output of nonideal normative theories, but that of *ideal* theories too.³

1. MOTIVATIONAL LIMITATIONS AND THE DEMANDS OF JUSTICE

Suppose we have a theory of justice that requires people to optimize their productive output while earning the same income as everyone else. Doctors should provide the best health care they can provide, manufacturers should produce the products people want at the desired quantity and quality, and janitorial staff should keep schools and offices optimally clean; and doctors, manufacturers, and janitors should all end up with the same net income.⁴ One might object to such a theory of justice by arguing that its fulfillment lies beyond what people can, in general, be motivated bring about. Since the costs of being a doctor are much higher than those of being a janitor (in terms of training hours, effort, and stress, for example), many of the highly skilled people we want to become doctors rather than janitors won't do so if the material rewards of being a doctor are roughly equal to those of being a janitor. Put simply, this theory of justice requires a level

2. This conclusion apparently engages a more recent paper by Estlund, in which he argues that that “the truth about justice is not constrained by considerations of the likelihood of success in realizing it” (David Estlund, “Utopophobia,” *Philosophy & Public Affairs* 42, no. 2 (2014): 113–134, at p. 115). If my argument in this paper is correct, then it turns out, *pace* Estlund, that “ought implies reasonably likely”, at least to the extent that justice does not require the realization of states of affairs that the relevant agents are too unlikely to realize due to their lack of motivational resources. As Estlund's more recent paper raises issues that are subtly distinct from those addressed here, I'm afraid I cannot engage it in greater depth without distracting from the argument in this paper

3. *Pace*, e.g., Ingrid Robeyns, “Ideal Theory in Theory and Practice,” *Social Theory and Practice* 34, no. 3 (2008): 341–362; A. John Simmons, “Ideal and Nonideal Theory,” *Philosophy & Public Affairs* 38, no. 1 (2010): 5–36; Zofia Stemplowska, “What's Ideal About Ideal Theory?” *Social Theory and Practice* 34, no. 3 (2008): 319–340; Laura Valentini, “On the Apparent Paradox of Ideal Theory,” *Journal of Political Philosophy* 17, no. 3 (2009): 332–355.

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

Motivational Limitations on the Demands of Justice

of self-sacrifice that is beyond human motivational limitations. The objection continues: If people can't be motivated to comply with the theory's requirements, then they can't comply with those requirements. If people can't comply with the theory's requirements, and if "ought" implies "can", then the proposed theory of justice is false — it is not the case that people ought to optimize their productive output while earning the same income as everyone else.

On its face, this line of reasoning seems plausible enough. "We want a theory of justice for humans, and such a theory must be sensitive to human motivational limitations", the proponent declares. Not every limitation, of course. For instance, transitory, perhaps even cultivated, selfishness or laziness should not be taken to block moral requirements. But, if there are any genuine motivational *incapacities* — limits on what people can be motivated to do that are hard-wired into human nature — then surely no plausible theory of justice can make demands that contravene them.⁵

There are three general strategies for denying this line of reasoning: the first denies that "ought" implies "can"; the second argues that the bounds of practical possibility ("can" in the relevant sense) are so malleable and indeterminate as to place few practical limitations on the demands of justice; the third denies that motivational incapacities delimit the relevant sense of "can".

Some might take G.A. Cohen to be a proponent of the first strategy (although this is potentially contentious). Cohen argues that ultimate normative principles are insensitive to (non-normative) facts; thus, the fundamental demands of justice are not limited by facts about what is feasible.⁶ It follows that what justice demands is insensitive to facts about which states of affairs we can be motivated to realize. Notice, though, that this strategy understands "ought" in an *evaluative* sense. Cohen's fact-insensitive principles tell us "not what we should do but what we should think, even when what we should think makes no practical difference."⁷ In other words, fact-insensitive principles are not meant to specify requirements for *action* (this is for "principles of regulation" to do); they are meant to direct our efforts to determine which states of affairs qualify as just or to rank

5. Not to say that there are motivational incapacities — or even such a thing as "human nature" — in this strongly essentialist sense. The point is just that, if there were motivational limitations of this sort, then a theory of justice would have to be constrained by *these* even if it weren't constrained by less essential limitations.

6. G. A. Cohen, *Rescuing Justice and Equality* (Cambridge, MA: Harvard University Press, 2008), ch. 6; cf. Anca Gheaus, "The Feasibility Constraint on the Concept of Justice," *The Philosophical Quarterly* 63, no. 252 (2013): 445–464; Andrew Mason, "Just Constraints," *British Journal of Political Science* 34 (2004): 251–268.

7. Cohen, *Rescuing Justice and Equality*, 268.

states according to some standard of moral desirability.⁸ This is a puzzling response to someone who thinks that the demands of justice are subject to feasibility constraints. Few have thought that “ ϕ is good” implies “can ϕ ”. The “ought” in “ought implies can” is generally understood as pertaining to the demands of morality upon *action*.⁹ This first strategy neglects this crucial presupposition.¹⁰

A second strategy might be thought more promising. This strategy concedes that the demands of justice upon action might be limited by facts about the states of affairs we are able to realize, but only by facts about “second-order” abilities, abilities that we may not possess now but that we can acquire by a protracted series of steps into the indefinite future. If a theory of justice demands that we realize a state of affairs that lies beyond the foreseeable feasibility frontier, then justice imposes upon us “dynamic duties” to bring about the conditions required to satisfy the theory’s requirements (including the expansion of our motivational capacities).¹¹ Hence, this strategy argues that the limits of our abilities are so malleable and indeterminate that “can” in the relevant sense places few (if any) limitations on what justice can properly demand with respect to action.

This second strategy fails to show that a valid theory of justice can make demands upon action that contravene human motivational limitations. Sure enough, it concedes that a directive is a proper demand of justice only if some agent(s) can acquire the necessary second-order abilities by a protracted series of steps into the indefinite future. But this leaves open the possibility that motivational deficiencies undermine the acquisition of second-order abilities. Perhaps I face a demand to give more of my income to charity but I am not motivated to do so. Suppose there is some series of steps that will instill in me the necessary motivational states, leading me to satisfy the demand to give more to charity. What if my motivational deficiencies are such that I fail to take some of the steps required to cultivate a motivation to comply with the demand? Do I have a second-order ability to give more to charity? For all the second strategy says, the answer might be no. If this is so, then I do not have a duty (dynamic or otherwise) to give more to charity. Thus, this strategy leaves open the possibility that the demands of justice upon action are constrained by human motivational limitations.

This is where the third strategy enters the fray. This strategy amounts to analyzing

8. For further discussion on the distinction between “fundamental principles” and “rules of regulation”, see Cohen, *Rescuing Justice and Equality*, pp. 265–268.

9. Cf. Peter B. M. Vranas, “I Ought, Therefore I Can,” *Philosophical Studies* 136 (2007): 167–216.

10. For further critical discussion, see David Wiens, “‘Going Evaluative’ to Save Justice From Feasibility — A Pyrrhic Victory,” *The Philosophical Quarterly* 64, no. 255 (2014): 301–307.

11. Pablo Gilabert, *From Global Poverty to Global Equality* (New York: Oxford University Press, 2012); Mark Jensen, “The Limits of Practical Possibility,” *Journal of Political Philosophy* 17, no. 2 (2009): 168–184.

Motivational Limitations on the Demands of Justice

the relevant sense of “can” so that the feasible set is not delimited by first-personal motivational limitations. That is, one might argue that the relevant sense of “can” is such that an agent is not permitted to take *her own* motivational limitations as marking the bounds of her abilities. An exemplar here is Holly Lawford-Smith’s analysis of feasibility in terms of an agent’s conditional probability of successfully realizing the target state of affairs given her attempt to do so.¹² First-personal motivational deficiencies might be thought to prevent attempts to realize a state of affairs; hence, by assuming trying, Lawford-Smith’s analysis excludes first-personal motivational limitations from the set of relevant feasibility constraints. Her rationale for doing so is plainly moral theoretic: “we don’t want to let agents off the moral hook. The fact that a person won’t do something isn’t enough for us to retract an imperative that she ought to.”¹³ In other words, if we agree that “ought” implies “can”, then our analysis of “can” should not permit people to excuse themselves from the demands of justice by (cynically) appealing to their own motivational limitations.¹⁴

This reasoning begs the question. To put a sharp point on it, we can reconstruct Lawford-Smith’s argument thus: given “ought implies can”, our analysis of “can ϕ ” should assume away the motivational deficiencies of the addressees of an obligation to ϕ —by assuming that they try to ϕ —because we do not want our analysis of “can ϕ ” to permit agents to avoid an obligation to ϕ by appealing to their motivational deficiencies. Put this way, the explanatory premise plainly assumes the claim to be shown: that barriers to trying, like motivational deficiencies, cannot block abilities and, in turn (by “ought implies can”), moral requirements.

Estlund’s argument might be thought a helpful supplement at this point. While conceding that “ought” implies “can”, Estlund argues that motivational deficiencies— even those that are hard-wired into human nature— do not undermine agents’ ability to satisfy the requirements of justice, even if motivational incapacities in fact prevent us from acting

12. Holly Lawford-Smith, “Understanding Political Feasibility,” *The Journal of Political Philosophy* 21, no. 3 (2013): 243–259; cf. Geoffrey Brennan and Nicholas Southwood, “Feasibility in Action and Attitude,” in *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz*, ed. T. Rønnow-Rasmussen et al., Available online at www.fil.lu.se/hommageawlodek. (2007).

13. Lawford-Smith, “Understanding Political Feasibility,” 254f.

14. Lawford-Smith takes a more nuanced position elsewhere, allowing an action A to qualify as “unavailable” for an agent if the agent’s evidence (e.g., a record of her past behavior or information about her personality) justifies a robust prediction that she will not A when presented the opportunity to do so (see Holly Lawford-Smith, “Non-Ideal Accessibility,” *Ethical Theory and Moral Practice* 16 (2013): 653–669, at p. 660). Thus, Lawford-Smith’s considered view might be that motivational deficiencies pose constraints on the demands of justice under certain conditions (cf. *ibid.*, pp. 667f). As my purpose here is not to sort out Lawford-Smith’s considered view, I set aside further discussion of this point.

in the required ways.¹⁵ As Estlund often puts it, “can’t will” does not imply “can’t do”, even if “can’t will” implies “won’t do”.¹⁶ Having the ability to, say, give a small portion of my income to charity is compatible with the fact that it is not possible that I be motivated to do so. So even if “ought” implies “can”, “ought” does not imply “can will”. It follows that motivational incapacities *never* block moral requirements. Thus, Estlund seems to have persuasively established that justice — even “justice-for-humans” — can make demands with which we cannot — perhaps even by human nature — be motivated to comply.

Any attempt to establish the general claim that human motivational deficiencies can, under certain circumstances, constrain the demands of justice must go through Estlund’s argument to the contrary. In this paper, I show not only that Estlund’s powerful counterargument is mistaken, but that the centerpiece of his argument — his analysis of “ability” — implies that motivational deficiencies constrain the demands of justice under at least one common condition. But before we get to the details of his argument, we should be clear about the sense of “ought” Estlund has in mind when he says that motivational incapacities are not requirement-blocking. Estlund’s distinction between “institutional proposals” and “institutional principles” is crucial here. The important difference boils down to this: an institutional *proposal* instructs us to implement a particular institutional scheme given how actual people are likely to act, while an institutional *principle* requires us to build a particular institutional scheme and then comply with it given how people can act.¹⁷ Accordingly, Estlund argues that the demands of our institutional *principles* are left unblocked by motivational incapacities. Take the theory of justice with which we started section 1, which Estlund calls “Pretax Max”. If someone were to propose that we implement the institutional scheme required to realize Pretax Max, then the fact that people are generally unlikely to act in the ways required by the institutional scheme would be sufficient to refute instructions to implement Pretax Max. But if someone claims that we should implement *and comply* with an institutional scheme that realizes Pretax Max — if Pretax Max is put forward as an institutional *principle* — then it is “*not*

15. Estlund, “Human Nature,” 228. The phrase “motivational incapacities” is shorthand for “motivational limitations hard-wired into human nature”. Again, neither Estlund nor I make any commitment to the existence of such things.

16. Although Estlund often suggests that the “can’t” in “can’t will” is the can’t of inability (as in “can’t do”), this is incorrect — Estlund does not wish to treat willing as a kind of doing. By “can’t will ϕ ”, Estlund means a restricted modal claim like the following: “it is not possible to will (i.e., be motivated to) ϕ given the relevant facts about the context of evaluation” (personal correspondence). This difference in the meaning of “can” should be read throughout.

17. Estlund, “Human Nature,” 217f.

Motivational Limitations on the Demands of Justice

refuted by any facts about whether people will build or comply with it.¹⁸ For the sake of generality, let's say that institutional proposals proffer *all-things-considered oughts* while institutional principles proffer *normative oughts* (for lack of more informative terms).¹⁹ The former direct agents how to act given how humans are likely to behave (thus ignoring unlikely possibilities), while the latter direct agents how to act given how humans are able to act. Estlund grants that motivational incapacities affect how we are likely to behave and, so, they constrain all-things-considered oughts. The central issue for the remainder of the paper is whether motivational incapacities delimit how we are able to act and, so, whether they constrain the set of valid normative oughts.²⁰

2. DOES "OUGHT" IMPLY "CAN WILL"?

Estlund says "no". How does he get to this conclusion? Estlund grants that "ought" implies "can", by which he means (as we've seen) that an alleged requirement of justice is refuted if there are no individual or group the ability to fulfill its demands.²¹ Importantly, "ought" does not imply "will do" — a theory of justice is not refuted if people will not, as a matter of fact, do what is necessary to meet its requirements.²² But is a theory of justice refuted if people can't be motivated to do what is necessary to meet its requirements?

Along with Estlund, we need not commit to the existence of genuine motivational incapacities. We are only considering whether characteristically human motivational incapacities, *if there are any*, can block normative oughts. So suppose there are some non-pathological motivational incapacities that are characteristic features of human psychology — they are "hard-wired", so to speak. Examples might be revulsion, fear, certain temptations, perhaps deep-rooted selfishness and partiality. We often talk as if we are unable to act a certain way when we are too afraid or too revolted to do so. "I

18. *ibid.*, 218, original emphasis; cf. 226.

19. Perhaps Estlund's language of "concessive" and "nonconcessive" oughts would work here (see, e.g., Estlund, "Utopophobia"). But this is not obvious to me, so I set those terms aside to avoid raising potential confusion.

20. By limiting "can" to "human abilities", I'm setting aside a discussion of feasibility constraints more generally. I present my own view on this issue in David Wiens, "Political Ideals and the Feasibility Frontier," *Economics and Philosophy* (forthcoming). For additional discussions, see Geoffrey Brennan and Philip Pettit, "The Feasibility Issue," in *Oxford Handbook of Contemporary Philosophy*, ed. Frank Jackson and Michael Smith (Oxford: Oxford University Press, 2005); Tyler Cowen, "The Importance of Defining the Feasible Set," *Economics and Philosophy* 23, no. 1 (2007): 1–14; Gilabert, *From Global Poverty*; Lawford-Smith, "Understanding Political Feasibility."

21. Estlund, "Human Nature," 210.

22. *ibid.*, 210f.

can't climb the ladder" is a natural thing to say when we are gripped by a (normal) fear of heights; "I can't eat that" is a normal reply when we are faced with the prospect of eating rotting, maggot-infested meat. There is a natural line of reasoning underlying this thought. Let ϕ be a generic action. If being motivated to ϕ is a precondition for doing ϕ and we can't be motivated to ϕ , then we can't satisfy a precondition for doing ϕ . Thus, we can't ϕ . Given this line of reasoning, "can't will to ϕ " implies "can't do ϕ ". So, according to this reasoning, if "ought ϕ " implies "can ϕ ", then "ought ϕ " implies "can will to ϕ ".

If Estlund wishes to insist that human motivational incapacities do not block moral requirements, he faces a choice: he must deny that "ought" implies "can" or he must deny that "can't will" implies "can't do".²³ He opts for the latter. To do this, he needs an analysis of "can" (in the sense of ability) that makes "can do" consistent with "can't will" (equivalently: "can't bring oneself to do"). The following seems to do the trick:

A person is able to (can) do something if and only if, were she to try and not give up, she would tend to succeed.²⁴

For Estlund, it seems motivational incapacities only prevent people from trying to ϕ or following through to ϕ 's successful realization. Simply put, "can't will" only implies "won't do". Since "won't do" is consistent with "can do", "can't will" does not imply "can't do". (Estlund's analysis of "ability" in the case of groups differs; I set this aside until section 3.)

Estlund illustrates this point with the case of "selfish Bill", in which Bill is so deeply selfish that he fails to comply with a requirement to not dump his trash by the road.²⁵ Estlund claims that it is outlandish for Bill to claim an exemption from the trash disposal requirement on account of his selfishness. Clearly, Bill is able to refrain from dumping his trash at the road—were he to try to dispose of his trash properly and not give up, he would tend to succeed. Bill is able to comply with the requirement; he simply fails to make a good faith effort to do so.

It is worth noting here that conditional analyses of "ability", like Estlund's, have fallen out of favor. Of particular relevance here is the fact that cases involving psychological shortcomings that prevent trying of the right sort are typically thought to be counterexamples to such analyses.²⁶ Indeed, one common response to selfish Bill expresses the

23. Estlund, "Human Nature," 213.

24. *ibid.*, 212.

25. *ibid.*, 220.

26. For useful overviews, see Bernard Berofsky, "Ifs, Cans, and Free Will: The Issues," in *The Oxford Handbook of Free Will*, ed. Robert Kane (New York: Oxford University Press, 2002); John Maier, "Abilities," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (2011), sec. 3.

point: “If Bill’s psychology is such that he really can’t bring himself to try to dispose of his trash properly — if Bill isn’t simply a jerk — then we should treat him as *unable* to dispose of his trash properly.” This is not the place to rehearse the problems facing conditional analyses of “ability”; nor is it the place to consider the merits of alternative analyses. This paper is centrally concerned with showing that, assuming “ought implies can”, motivational deficiencies can constrain the demands of justice. To my knowledge, Estlund’s is the most forceful argument denying this general claim. A survey of the wider literature on “ability” might provide more comprehensive support for the claim that motivational deficiencies can be ability-blocking; but it would be a less dialectically effective strategy in the shadow of Estlund’s argument. After all, Estlund (and his sympathizers) might simply deny the alleged counterexamples to his analysis of “ability”.²⁷ Rather than mount a comprehensive assault on Estlund’s analysis of “ability”, I grant it for dialectical purposes and show that justice can be constrained by motivational deficiencies on the most plausible specification of the key phrase “try and not give up”.

3. “CAN’T WILL (IN GOOD FAITH)” IMPLIES “CAN’T DO”

In this section, I show that Estlund’s argument rests on an ambiguous analysis of “ability”, the most plausible specification of which implies that motivational incapacities undermine a person’s ability to act under at least one common condition — namely, so long as the agent makes a sincere effort to perform the action in question in the face of her motivational deficiencies. For convenience, let’s say that an agent is subject to a “good faith motivational incapacity” when this condition is satisfied. Put simply, my central claim is this: for good faith motivational incapacities, “can’t will” implies “can’t do”. Thus, *pace* Estlund, “ought” implies “can will (in good faith)” if “ought” implies “can”.

Recall Estlund’s analysis of “can”: “A person is able to (can) do something if and only if, were she to try and not give up, she would tend to succeed.” What does it mean for a person to “try and not give up”? (I take up group abilities in the next section.) When has a person satisfied this crucial condition? Let’s start by considering two relatively implausible candidate specifications for the sake of introducing key parameters for further refinement.

- (1) A person is able to (can) ϕ if and only if, were she to complete a sequence of acts that conduces to ϕ , she would tend to ϕ successfully.

27. Many of these counterexamples involve *pathological* psychological limitations; see, for instance, Lehrer’s classic case of the red sugar balls in Keith Lehrer, “Cans Without Ifs,” *Analysis* 29, no. 1 (1968): 29–32. But Estlund concedes that pathological motivational incapacities can be requirement-blocking. So cases such as these will not worry him much.

This can't be right: sometimes people are unable to ϕ precisely because they are unable to complete a sequence of acts that conduces to ϕ , even if they are able to start such a sequence. For instance, many people cannot do a reverse two-and-a-half somersault in pike position from a 10 meter diving platform. For at least some of these people, this is because they cannot do a backflip, which is a constituent of a reverse two-and-a-half somersault in pike position. Since they cannot do a backflip, they cannot complete a sequence of acts that has a reverse two-and-a-half somersault in pike position as a possible outcome.²⁸ Yet, according to (1), these people might nonetheless be able to perform the dive, since they might tend to perform it successfully if they were to complete the required sequence of acts, including the backflip. More generally, (1) is an implausible specification of Estlund's analysis because it makes an ability to ϕ compatible with an inability to complete a sequences of acts required to make ϕ possible. Note that we can't remedy this defect by saying that a person can ϕ if and only if she is able to complete a sequence of acts that conduces to ϕ and, were she to do so, she would tend to ϕ . This simply analyzes an ability to ϕ in terms of an ability to complete a sequence that conduces to ϕ , which is uninformative.

Let's move on to the second candidate:

- (2) A person is able to (can) ϕ if and only if, were she to repeatedly initiate a sequence of acts that conduces to ϕ , she would tend to ϕ successfully.

This remedies the defects of (1) by excluding people who are unable to complete a sequence that conduces to ϕ from the set of people who can ϕ . But it does so at the cost of stripping the ability to ϕ from people who initiate a sequence that conduces to ϕ but simply decline to complete it. Estlund's selfish Bill case can be modified to illustrate the point. Suppose that, each week, Bill starts out to bring his trash to the community disposal bin but, halfway there, he decides he can't be bothered to walk all the way to the bin and simply dumps his trash at the road. According to (2), Bill lacks the ability to dispose of his trash properly because he tends to dump his trash at the road despite repeatedly attempting to walk his trash to the bin. But, by assumption, Bill is simply being a jerk. Surely the fact that he can't be bothered to give more than a half-hearted effort does not undermine his ability to dispose of his trash properly. The important point here is that our analysis of "able to ϕ " should account for the sincerity (or lack thereof) with which people attempt to complete a sequence that conduces to ϕ .

The defects of (1) and (2) are easily fixed:

28. Note that "conduce" here must mean something like "has ϕ as a possible outcome" rather than "makes ϕ more likely", since the latter implies a tendency to ϕ successfully.

Motivational Limitations on the Demands of Justice

- (3) A person is able to (can) ϕ if and only if, were she to repeatedly make good faith attempts to complete a sequence of acts that conduces to ϕ , she would tend to ϕ successfully.

Whereas (1) includes too many people in the set of people who can ϕ , and (2) includes too few, (3) seems about right as a specification of Estlund's original analysis — it does not confer an ability to ϕ to people who are unable to complete a conducive sequence nor does it take that ability away from those who initiate and subsequently decline (in bad faith) to complete a conducive sequence. Importantly, though, (3) permits motivational incapacities to undermine abilities in certain circumstances.

For instance, writing (and completing) a book requires certain motivational capacities — enough resilience to press through inevitable challenges that threaten to halt the enterprise, enough stamina to press through bouts of despair and distraction, enough confidence to absorb criticism, and so on. Now suppose both Danny and Claudia attempt to write a book. Danny produces a few pages but realizes after a week that writing a book is harder than he anticipated and that he can't be bothered to follow through on his initial intention. He abandons his plan of writing a book and never gives it a second thought. Can Danny write a book? Maybe — were he to repeatedly make a good faith effort, perhaps he would successfully complete a book. According to (3), he would then qualify as being able to write a book. Simply giving up on writing a book is insufficient to block his having the ability to do so. Claudia also realizes after a week that writing a book is harder than she anticipated; but instead of giving up, she presses on the following week, producing a few more pages. Claudia has some good weeks, but she also has weeks where she suffers significant setbacks. After a year, Claudia's despair and lack of self-confidence get the best of her and she abandons her book. A couple of years later, Claudia has a new idea for a book, which she sets out to write. Again, after many weeks of encountering obstacles, despair and low confidence set in again and she abandons her book. This pattern repeats itself several more times, until Claudia finally gives up her dream of writing a book.

Can Claudia write a book? It seems she can't. Of course, Claudia can write down words, form sentences, and string sentences together into cohesive paragraphs and chapters. Let's even suppose that her prose is impeccable. Yet she fails to complete a sequence of acts that conduces to completing a book, despite repeatedly making good faith efforts to do so. According to (3), Claudia is unable to write (and complete) a book. This seems right — there's little reason to think that her ability to write impeccable prose translates to an ability to write a book (as there is little reason to think that an ability to run a mile translates to an ability to run a marathon). By assumption, Claudia's repeated failure to complete a book stems from her repeated failure to press through periods of

deep despair, her fear of rejection, and so forth. Put simply, Claudia's inability derives from a lack of important motivational resources.

Claudia's case is not unique. There are many cases that are structurally similar to hers, where achieving some goal requires carrying out a protracted, complex sequence of acts that is fraught with obstacles—becoming a doctor, building a dream home, composing a symphony, proving a mathematical theorem. Achieving goals like this requires certain motivational capacities to continue with the process despite palpable setbacks and uncertain success. According to (3), certain motivational deficiencies can undermine an ability to achieve the target goal. This seems right—an ability to perform different component acts does not always translate into an ability to put these component acts together in a protracted and complex sequence that conduces to achieving the goal or to sustain such a sequence over time. Thus, (3) permits some (non-pathological) motivational incapacities to undermine abilities under certain conditions. Specifically, (3) implies that an ability to ϕ is undermined by a lack of motivational resources that blocks the performance of ϕ despite sincere efforts to overcome the offending motivational deficiencies. Call these *good faith motivational incapacities*. My claim is this: if (3) is the most plausible specification of Estlund's analysis of "can" (it seems it is), then good faith motivational incapacities undermine abilities. As a slogan, "can do" implies "can will (in good faith)".²⁹

4. GROUP ABILITIES AND INSTITUTIONAL PRINCIPLES

Estlund's argument largely focuses on directives to build and comply with certain institutional schemes, which are typically out of reach for individuals taken in isolation. Not so for groups. So—can good faith motivational incapacities undermine a group's ability to build and comply with a morally desirable institutional scheme?

To fix ideas, let's consider a requirement to implement a constitutional democracy and subsequently comply with the enacted rules (I leave the "constitutional" qualifier implicit hereafter). For Estlund, a group's abilities are constituted by its members' abilities: a group of people *is able to* (or "can" in the relevant sense) do something only if the relevant members of the group are to able carry out the individual-level actions necessary to do

29. I don't claim that (3) is fully adequate as an analysis of "ability". Rather, I claim that (3) is the most plausible specification of Estlund's original analysis, which I grant here for the sake of argument. I conjecture that any counterexamples to (3) are counterexamples for any conditional analysis like Estlund's. Thanks to John Maier for pressing me to clarify this point.

Motivational Limitations on the Demands of Justice

that thing.³⁰ (The thing to be done might require coordinated action; if that's so, the set of necessary individual-level actions includes efforts to coordinate individuals' activity in the appropriate way. This proviso should be read throughout.) Hence, a society can build and comply with a democratic institutional scheme only if the relevant members of society are able to carry out the individual-level actions necessary to build a democratic institutional scheme and subsequently comply with its constitutive rules. By "ought implies can", this requirement is blocked if the relevant members of a society are unable to carry out the individual-level actions necessary to meet its demands. The question, then, is whether there are any good faith motivational incapacities that prevent people from carrying out the individual-level actions (including efforts to coordinate individual-level activity) necessary to build a democratic institutional scheme or subsequently complying with its constitutive rules.

Suppose a society is ruled by a pathologically paranoid and ruthless dictator, who meets dissent with vicious attacks against all dissidents and their family members. Subjects are known to snitch on fellows, even close relatives, who express dissent in private. Not even top military generals or state bureaucrats are safe. All members of society are given to silence for fear of reprisals against themselves and their family. Under conditions like this, society can build a democratic institutional scheme only if a sufficient number of people are able to organize and collectively resist the dictator's rule until the dictator steps down or is forcibly removed from power. In the past, very small groups of people (including military personnel) have tried to organize resistance campaigns, but they have always been put down effectively. However, unbeknownst to the dictator's subjects, the dictator's coercive capacities are quite limited: if as little as two percent of the population were to participate in a moderately sustained resistance campaign, the dictator would cede his power and a path to constitutional democracy would open. (For illustrative purposes, the two percent figure amounts to roughly 500 000 North Koreans or 1.5 million Iranians.)

At a glance, we might judge that this society is able to build a constitutional democracy. All that's required is for a relatively small proportion of the population to organize and join sustained public protests in locations throughout the country. At the individual level, this requires a person to enter a public square for some number of days to express opposition

30. Estlund, "Human Nature," 236. Collective *agents* are typically thought to satisfy more stringent conditions than mere aggregates; see, e.g., Holly Lawford-Smith, "The Feasibility of Collectives' Actions," *Australasian Journal of Philosophy* 90, no. 3 (2012): 453–467 and citations therein. Following Estlund, I leave aside the issue of collective agency and focus on the abilities of groups as such (of which collective agents are a subset).

to the dictator and to persuade others to join her in protest. If a small percentage of the population does this simultaneously, the dictator will step down. There are potential external barriers to the initial formation of the protest, of course; if the protest starts off too small, it will be put down before it is able to gain momentum. But these barriers disappear once the opposition movement is large enough—which simply means that enough individuals have entered the public square to participate in a coordinated protest. It seems plausible that at least two percent of the population is able to enter a public square and express opposition to the dictator.³¹

This judgment is too hasty. Consider the situation from the perspective of Maggie, an adult subject who is representative of the two percent who are most committed to overthrowing the dictator. In general, Maggie has shown great resolve in following through on fairly demanding commitments; in particular, she is willing to bear a fairly high risk of making important sacrifices to unseat the dictator. However, from her perspective, the expected cost of the required individual-level behavior is far beyond this threshold. All her evidence indicates that participating in public protest means that she will certainly sacrifice her life, the lives of her spouse and children, and probably the lives of several close friends. Further, all Maggie's evidence tells her that there is a good chance that any number of her close relatives and friends are government informants. Thus, even private intimations that she is considering organizing and joining a public protest carry a significant risk of sacrificing her life and the lives of those closest to her. Finally, Maggie safely assumes that most people have made roughly the same cost-benefit calculations on the basis of similar evidence (although she can't confirm this, for fear of being found out). Thus, Maggie estimates that the number of people who are willing to engage in public protest is too low and, accordingly, that the likelihood of a successful opposition movement is correspondingly low. In sum, Maggie (and others like her) expects that, were she to suggest to anyone that she might oppose the dictator, she would most likely pay extreme costs with very little prospect of successfully ousting the dictator.³²

Now suppose Maggie resolves to tell her spouse that she opposes the dictator and

31. One might assert that the society is able to build a constitutional democracy if the dictator voluntarily steps down or ceases to repress dissent and the dictator is able to step down or cease repression. I've tried to circumvent this response by assuming that the dictator is pathologically paranoid and ruthless, hoping to pump the intuition that he really can't step down or cease repression. If this is not convincing, then I limit myself to inquiring whether the populace is able to build (and comply with) a democratic institutional scheme holding the dictator's conduct fixed. This seems like an appropriate limitation, since the dictator's conduct is not subject to citizens' control.

32. Cf. Timur Kuran, "Now Out of Never: The Element of Surprise in the East European Revolution of 1989," *World Politics* 44, no. 1 (1991): 7–48.

Motivational Limitations on the Demands of Justice

wants to help organize a public protest. As she sits across the table from him, she becomes anxious about the extreme risk she is taking — “What if he is an informant? What if my children are sentenced to a labor camp?” The risk of telling her husband of her intention becomes too much to bear and she suppresses her subversive thoughts. After a while, she regains some courage and resolves to tell her closest friend about her desire to organize a public protest. Again, as she sits across the table from her friend, she becomes overwhelmed by the risks she is about to incur and she remains silent. The same story repeats itself over several years. Each time, despite her repeated good faith efforts to overcome her fear, Maggie cannot bring herself to reveal her opposition to the dictator. The risks are simply too great for her to bear. Maggie can’t, in good faith, “muster the will” to express opposition to the dictator or persuade others to join an opposition movement.

Given (3), Maggie is unable to organize and join a sustained opposition campaign because, despite her repeated good faith efforts to complete a sequence of acts that conduces to organizing and joining an opposition campaign, she does not successfully do so. Maggie’s inability to organize and join an opposition campaign is due to various motivational incapacities — her risk aversion and her insufficient courage, among others.³³ Recall that the society is able to implement democratic institutions only if enough people participate in a sustained campaign to overthrow the dictator. Since Maggie is typical of the segment of the population that is most deeply committed to overthrowing the dictator, there are too few people who are able to coordinate and participate the sort of opposition campaign required to pave the way for democratization.³⁴ So this is a case where individuals’ good faith motivational limitations undermine society’s ability to build (and, thus, build and comply with) a democratic institutional scheme.

Several objections arise here. First, one might argue that Maggie makes a calculated decision not to express dissent in light of the risks, not that she can’t be motivated to express dissent (given the relevant contextual specifics). But this diagnosis misunderstands the case. The case presupposes that, when coolly comparing the costs and benefits of expressing dissent versus remaining silent, she repeatedly forms the intention to express

33. I leave open the possibility that Maggie might acquire an ability to organize and join an opposition movement if she were to revise her estimation of the risks involved in light of new evidence.

34. This is not to say that there aren’t a few individuals — the zealots — who are less risk-averse or more courageous (or foolhardy) than Maggie. The point is that, by assumption, the zealots do not form the two percent required for a successful resistance campaign. In other words, given that Maggie is typical of the two percent of people who are *most* committed to overthrowing the dictator (where commitment is indicated by some measure of the risk one is willing to bear), I’m assuming for illustrative purposes that at least two percent of the populace isn’t sufficiently risk-acceptant or courageous to instigate a successful resistance movement.

dissent. That is, she repeatedly decides to express dissent even in view of an accurate assessment of the risks. Moreover, she initiates courses of action that aim to culminate in an expression of dissent — she enters meetings with potential confidants sincerely intending to reveal her opposition to the dictator. What prevents her from ultimately expressing dissent is the fear that grips her at the climactic moment.³⁵

Alternatively, one might object that Maggie is not really unable to participate in an opposition movement; rather, she has the relevant ability but her use of it is thwarted by the climate of fear created by the dictator.³⁶ To wit, we know that at least a few people have publicly resisted the dictator's rule — namely, those who have served as examples of the dictator's ruthless treatment of dissidents. We might suppose that Maggie is like these actual dissidents in all respects pertaining to their abilities. Since the punished dissidents had the ability to publicly oppose the dictator, so Maggie (and others like her) must have that ability too; unlike the punished dissidents, Maggie is simply too afraid to exercise the ability. This reply simply assumes that (repeatedly) succumbing to fear cannot undermine one's abilities. If the punished dissidents did have the ability to publicly oppose the dictator,³⁷ then, according to (3), had they made repeated attempts to publicly express opposition, they would have tended to succeed. But this is not true of Maggie: she fails to publicly express opposition to the dictator despite repeatedly making good faith attempts to complete a sequence of acts that conduce to publicly expressing opposition to the dictator. So Maggie is unlike the punished dissidents in at least one respect pertaining to their abilities.

Sustaining this objection requires us to revise (3), since Maggie is clearly unable to publicly oppose the dictator according to that analysis. In the spirit of the objection, we might deploy a notion of “normal circumstances” as follows: A person is able to (can) ϕ if and only if, were she, *under normal circumstances*, to repeatedly make good faith attempts to complete a sequence of acts that conduces to ϕ , she would tend to ϕ successfully. We might go on to say that the climate of fear fostered by the dictator does not qualify as normal in the relevant sense; under normal (i.e., less repressive) circumstances, Maggie would tend to publicly oppose the dictator given her repeated good faith attempts to do so. But now we need an analysis of “normal” that does not simply presuppose

35. Thanks to Holly Lawford-Smith for pressing me to clarify this point.

36. This objection is due to an anonymous referee.

37. It's not so clear that the actual dissidents did have this ability — actual performance does not imply ability (cf. Estlund, “Human Nature,” 212). For example, I might, by a stroke of luck, shoot a basketball in the hoop from half court on my first attempt. This does not imply that I have the ability to sink a half court shot. The analysis in (3) captures this: I would be unable to sink a half court shot if I tended to miss the shot despite repeated good faith attempts to make it. I set aside this complication hereafter.

Motivational Limitations on the Demands of Justice

that good faith motivational incapacities are not ability-blocking. In general, “under normal circumstances” cannot simply mean “in the absence of the environmental factors that trigger the motivational limitation at issue”; in this specific case, “under normal circumstances” cannot simply mean “in the absence of the fear-inducing features of Maggie’s environment”. Then the suggested revision of (3) would be plainly ad hoc. We need case-independent reasons to exclude fear-inducing circumstances from the set of “normal circumstances”; there don’t seem to be any good reasons to do so, as there is nothing especially unusual about fear-inducing circumstances in general. Moreover, the sort of environment described in the case seems to qualify as “normal” for the class of circumstances under which individuals need to develop and exercise the ability to publicly oppose dictators. Think of Stalinist Russia, communist East Germany, Cambodia under the Khmer Rouge, or contemporary North Korea; these are not unusual cases of totalitarian regimes, but prime exemplars. It is difficult to see how we might revise (3) in a principled way so that Maggie in fact has the ability to publicly oppose the dictator.

Estlund suggests a third objection, claiming that we might deny a requirement to build and comply with a democratic institutional scheme in this case without conceding that the content of moral requirements is constrained by human motivational incapacities. We start by noting that Maggie’s risk aversion is “morally sound or reasonable” as a matter of “substantive moral judgment about [its] content”. We then deny a requirement to bear such significant risk on the basis of this substantive moral judgment.³⁸ Thus, we have good moral reasons to deny a requirement to build and comply with a democratic institutional scheme in the above case, apart from any facts about human motivational limitations.

This reply is a red herring. The key issue here is whether, assuming “ought” implies “can”, there are any motivational incapacities that are sufficient to block moral requirements. Perhaps it is true that we need not appeal to facts about human motivational limitations to deny a requirement to bear heavy risks in the service of fundamental political reform. But this only shows that appeal to facts about motivational incapacity are not *necessary* to block the requirement; it doesn’t show that motivational incapacities are not *sufficient* to block the requirement. Perhaps our denial of the requirement is overdetermined. This is compatible with the claim that good faith motivational incapacities are sufficient to block moral requirements. I take it that Estlund set out to deny this sufficiency claim, not the necessity claim.

Estlund would reply that this leaves us with the following sort of argument against a

38. *ibid.*, 223. Here, I am extrapolating from Estlund’s strategy for dealing with humans’ alleged natural partiality and any motivational incapacity that might derive from this; see *ibid.*, sec. VI.

requirement to bear extreme risks in the service of political reform:

This tendency to eschew extreme risks in the service of political reform is what people are motivationally like, as a matter of human nature. *Therefore*, requirements to do otherwise are specious and false.³⁹

But this is a problematic sort of argument because it apparently licenses the following structurally similar argument:

People tend to a certain degree of cruelty, and this is part of what they are motivationally like as a matter of human nature. (Suppose this is so.) *Therefore*, requirements to be otherwise are specious and false.⁴⁰

Let's grant to Estlund that the latter is an "absurd" argument. Yet we need not deny the propriety of the former argument. In light of (3), the former argument is an instance of the following general argument:

People show a tendency to some good faith motivational incapacity, as a matter of human nature. *Therefore*, requirements that contravene this good faith motivational incapacity are specious and false.

If aversion to extreme risks is a good faith motivational incapacity and cruelty is not, then the first argument does not license the second one, their structural similarity notwithstanding. We can accept the first argument without accepting the second and, thus, maintain our claim that good faith motivational incapacities are sufficient to block moral requirements.⁴¹

In sum, a society is able to satisfy an institutional principle only if the relevant members of that society are able to carry out the individual-level actions required to bring

39. Cf. Estlund, "Human Nature," 224.

40. *ibid.*

41. If an impulse to be cruel or selfish turns out to be a good faith motivational incapacity—that is, if people typically do not refrain from engaging in cruel or selfish acts despite repeated sincere efforts to overcome the relevant motivational deficiencies—then we might adopt Peter Graham's proposal that "positive" duties—duties to perform some ϕ —are subject to the "ought implies can" principle, while "negative" duties—duties to refrain from performing some ϕ —are not (see Peter A. Graham, "'Ought' and Ability," *Philosophical Review* 120, no. 3 [2011]: 337–382). Given this proposal, the first argument in the text does not license the second argument because the first pertains to a positive duty—a requirement to undertake some political reform—while the latter pertains to a negative duty—a requirement to refrain from cruelty. I'm indebted to Holly Lawford-Smith for suggesting this point.

Motivational Limitations on the Demands of Justice

about the collective action needed to satisfy the institutional principle. Since Maggie is representative of those in society who are most committed to overthrowing the dictator, it turns out that too few people are able to coordinate and participate in an opposition campaign. Thus, Maggie's society is unable to build a democratic institutional structure. Crucially, this inability is largely due to widespread motivational limitations within the society. Since this society cannot build a democratic institutional scheme, it cannot build and comply with a democratic institutional scheme. Thus, good faith motivational incapacities can undermine a society's ability to build and comply with a morally desirable institutional scheme. Given "ought" implies "can" (and Estlund's analysis of "can"), it follows that good faith motivational incapacities can block a requirement to build and comply with a morally desirable institutional scheme.⁴²

5. "UNLIKELY TO WILL (IN GOOD FAITH)" IMPLIES "CAN'T DO"

Perhaps Estlund will be content to have refuted the strong claim that motivational incapacities *always* constrain the content of justice and accept the weaker claim that motivational incapacities can *sometimes* constrain the content of justice (although he claims to refute the weak claim too⁴³). Yet my argument implies something stronger than the claim that good faith motivational incapacities block moral requirements. In this section, I show that, on the most plausible reading of Estlund's analysis of "ability", good faith motivational *improbabilities* block moral requirements too. In stark contrast with Estlund's stated aim, his analysis of "ability" (adequately specified) implies that the content of justice depends, to some extent, on how people are *sufficiently likely* to be motivated. It follows, contrary to conventional methodological wisdom, that the distribution of good faith motivational capacities we find in the actual world constrains not only the output of nonideal normative theories, but that of *ideal* theories too.

To start, note that, given Estlund's analysis (on any interpretation), successfully ϕ -ing is not sufficient to have an ability to ϕ . To be able to ϕ , one must *tend* to ϕ successfully given repeated good faith attempts to do so. Using Estlund's own example, the mere fact that I "draw a jack of hearts from a shuffled deck in a single draw" doesn't show that I have

42. To be clear, everything I have said is consistent with the claim that a constitutional democracy is morally superior to brutal dictatorship — that a world in which this society builds and complies with a democratic institutional scheme is ranked above the status quo according to some moral evaluative standard. If one must translate this thought into the language of "requirements" or "oughts", then I concede that we ought to build and comply with a democratic institutional scheme in this evaluative sense of "ought". But recall that the issue is whether good faith motivational incapacities are sufficient to block requirements for action.

43. Estlund, "Human Nature," 228.

an ability to do so; I must also tend to do so given a large enough sample of good faith attempts.⁴⁴ I'm not sure what, precisely, "tend" means here, but I assume it must imply something like "sufficiently likely".⁴⁵ It follows that I am able to draw a jack of hearts from a shuffled deck only if, given repeated good faith attempts to do so, the likelihood of my drawing a jack of hearts is sufficiently high.

At a glance, abilities to realize protracted one-off achievements like writing a book or becoming a doctor seem to pose a problem for this alleged connection between ability and a tendency to succeed. Surely Claudia is able to write a book even if she writes (and completes) a book solely on the 17th of 25 attempts. However, as with Estlund's jack of hearts case, it is implausible to assert that Claudia is able to write a book merely because there exists a possible world at which Claudia completes a book. This thought suggests a resolution: we can (and should) preserve a connection between "ability" and some notion of "tendency to succeed" by looking at patterns across modal space. For example, we should say that Claudia has a tendency to successfully complete a book if and only if her completing a book is a "likely enough" outcome, in the sense that she successfully completes a book in a high enough proportion of the possible worlds⁴⁶ at which she repeatedly makes good faith attempts to complete a sequence of acts that conduces to completing a book. For purely illustrative purposes, if Claudia completes a book on 1 out of every (e.g.) 25 attempts in 75% of the worlds at which she makes repeated good faith efforts to write a book, then we might say that Claudia has a tendency to complete a book in the relevant sense — it is a likely enough possibility. If this were true, then (3) implies that Claudia is able to complete a book. Alternatively, if Claudia completes a book on (e.g.) 15 out of every 25 attempts in 5% of the worlds at which she makes repeated good faith attempts and never completes a book at any of the remaining 95% of the relevant worlds, then we might say she lacks a tendency to complete a book in the relevant sense — it is a highly unlikely outcome. In this case, (3) implies that Claudia is unable to write a book. Both of these judgments seem correct (although one should not get hung up by the particular numbers used to illustrate the point). More generally, we can refine (3) as follows:

- (4) A person is able to (can) ϕ if and only if she successfully completes a sequence of acts that manifests ϕ in a sufficiently high proportion of the possible worlds at

44. Estlund, "Human Nature," 212.

45. I don't think anything turns on how we understand probabilities here, whether it be in terms of frequencies or dispositions or some other notion.

46. Understood, in the usual way, as a complete way the world might be.

Motivational Limitations on the Demands of Justice

which she repeatedly makes a good faith attempt to complete a sequence of acts that conduces to ϕ .⁴⁷

In section 3, we saw that a person can be unable to ϕ due to a good faith motivational incapacity to overcome obstacles to the completion of a sequence that conduces to ϕ . From (4), it follows that a person is unable to ϕ if she is sufficiently unlikely to have the motivational resources needed to overcome obstacles to completing a sequence of acts that conduces to ϕ . Thus, Claudia is unable to complete a book if she is sufficiently unlikely to have the motivational resources needed to overcome obstacles to completing a book despite her good faith effort to do so. Why? Because if this is the case, then she fails to complete a book in a high enough percentage of the possible worlds at which she makes repeated good faith efforts to do so. In other words, if Claudia is sufficiently unlikely to have the motivational resources to overcome obstacles to completing a book, then she tends to fail at writing a book despite her repeated good faith attempts to do so.

A similar thought applies to Maggie. If Maggie is sufficiently unlikely to have the motivational resources required to bear the extreme risks associated with opposing the dictator (e.g., the motivational resources to overcome her fear), then she overcomes the obstacles to organizing and joining an opposition campaign in a very low proportion of the possible worlds at which she repeatedly makes good faith attempts to do so. It follows from (4) that Maggie is unable to organize and join an opposition campaign against the dictator. Since Maggie is typical of a subset of the population that is most committed to overthrowing the dictator, too few members of society are able to coordinate and participate in an opposition campaign and, in turn, society is unable to build (and comply with) a democratic institutional scheme. If “ought” implies “can”, it follows that good faith motivational *improbabilities* can block a requirement to build and comply with a democratic institutional scheme.

In sum, given (4) as a further refinement of Estlund’s analysis of “can”, a person is able to ϕ only if she is sufficiently likely to have the motivational resources required to overcome obstacles to completing a good faith attempt to ϕ . As a slogan, “can do” implies “likely enough to will (in good faith)”. If “ought” implies “can”, it follows that “ought” implies “likely enough to will (in good faith)”.

47. As Maier notes, an ability to ϕ must bear some connection to actual performance of ϕ (Maier, “Abilities,” sec. 2.2). Accordingly, the actual world must be one of the worlds at which she successfully ϕ s. Since (4) is already a mouthful, I leave this caveat implicit.

6. “OUGHT” IMPLIES “LIKELY ENOUGH TO WILL (IN GOOD FAITH)”

In sections 3 and 4, I argued that “can do” implies “can will (in good faith)” if we accept Estlund’s analysis of “can” (as specified by (3)). In section 5, I argued that “can do” also implies “likely enough to will (in good faith)” if we accept (4) (as a further refinement of (3)). Thus, if we grant that “ought” implies “can”, then not only does “ought” imply “can will (in good faith)”; “ought” also implies “likely enough to will (in good faith)”. It remains to be seen what “likely enough” amounts to, but we need not settle this here. Nor need we enumerate the precise conditions under which motivational incapacities qualify as “good faith” motivational incapacities. I’ve shown, *pace* Estlund, that good faith motivational incapacities (indeed, improbabilities) — *if there are any such things* — are sufficient to block moral requirements. Thus, motivational deficiencies can be requirement-blocking under at least one condition — namely, when they block performance of the required action despite repeated sincere efforts to overcome the salient motivational limitations.

This conclusion has potentially far-reaching implications for normative political philosophy. Perhaps few philosophers share Estlund’s view that facts about human motivational limitations are wholly irrelevant for sorting out the requirements of justice. Yet conventional wisdom holds that so-called ideal theories of justice can (and should) idealize human agents by abstracting away from many (if not most) motivational deficiencies. Hence, ideal theories assume a human motivational structure that is favorable for the realization of justice; among other things, agents are assumed to have a “sense of justice” and a general willingness to comply with the demands of justice.⁴⁸ This is consistent with acknowledging that there are limits on possible human motivational capacities; but the limits conventionally acknowledged are beyond the limits set by the motivational capacities actual people are likely to have.⁴⁹ According to conventional wisdom, it might be the business of nonideal theory to work within the motivational limitations actual people are likely to have, but the demands of justice should not concede to the distribution of motivational capacities we find in the actual world. The argument of this paper challenges this thought. If we accept that the practical requirements of justice — requirements to “build and comply” with just institutions — are constrained by what people are able to do, then it is the business of ideal theory, too, to work with the motivational capacities people are likely enough to have.

48. Robeyns, “Ideal Theory in Theory and Practice”; Simmons, “Ideal and Nonideal Theory”; Stemplowska, “What’s Ideal About Ideal Theory?”; Valentini, “Apparent Paradox.”

49. Gilibert, *From Global Poverty*; Jensen, “The Limits of Practical Possibility.”