

Ideal Theory and ‘Ought Implies Can’

Amy Berg¹

Pacific Philosophical Quarterly (2018)

When we can't live up to the ultimate standards of morality, how can moral theory give us guidance? We can distinguish between ideal and non-ideal theory to see that there are different versions of the voluntarist constraint, 'ought implies can.' Ideal moral theory identifies the best standard, so its demands are constrained by one version. Non-ideal theory tells us what to do given our psychological and motivational shortcomings and so is constrained by others. Moral theory can now both provide an ultimate standard and give us guidance; this view also gives us new insights into demandingness and blame.

Introduction

There is an apparent tension between two tasks of moral theory. On the one hand, morality should provide a standard for us to live up to, a standard that does not yield to facts about our individual psychological or motivational shortcomings. But

¹ I received helpful feedback on this paper from Craig Agule, Richard Arneson, Saba Bazargan-Forward, Matt Braich, Jeffrey Brand, David Brink, Brookes Brown, Colin Chamberlain, Barry Maguire, Alex Rajczi, Geoff Sayre-McCord, Steve Swartzter, and Alex Worsnip. Previous versions of this paper were presented to audiences at Temple University, the University of Indianapolis, Claremont McKenna College, UC San Diego, the Pacific Division meeting of the APA, the Rocky Mountain Ethics Congress, and a conference on Ideals and Reality in Social Ethics at the University of South Wales. I am grateful to these audiences for their feedback. Finally, I'm grateful for the helpful comments provided by two anonymous referees and the area editor of *Pacific Philosophical Quarterly*.

at the same time, morality has to provide action guidance for us as we are, with our flaws and idiosyncrasies. How can morality guide our actions when we cannot live up to its standard?

This tension is related to another debate, about the role, plausibility, and correct interpretation of the *voluntarist constraint*, that *ought implies can*. If ought implies can, as is commonly believed, we cannot be obligated to do something that is impossible for us. But there is disagreement about what the constraint means. How strictly should we interpret the ‘can’ in ‘ought implies can’?

We can use the distinction between *ideal and non-ideal theory* to make progress on these two issues. First, I show why the tasks of moral theory are often incompatible. I then turn to the voluntarist constraint. I argue that there are different modalities appropriate to different sets of obligations and, hence, different readings of the voluntarist constraint. That is, there is no single ‘can,’ or ‘ought,’ in ‘ought implies can.’ We should recognize the ways in which different kinds of inability constrain different kinds of obligations for an agent.

Morality can both provide an ultimate standard and guide our actions when we make use of these different interpretations of the voluntarist constraint. This is where the distinction between ideal and non-ideal theory can help us.¹ Ideal morality identifies the best standard, without yielding to our flaws, and so its demands are constrained only by a thin version of the voluntarist constraint, using physical possibility. In contrast, non-ideal morality tells us what to do given our present situation and shortcomings. Its demands are constrained by thicker versions of the voluntarist constraint, taking into account features of our psychology and motivational structure. As we transition from non-ideal to ideal morality, we thin out the voluntarist constraint. When we use this approach, morality can accomplish all of its tasks. I close by showing how this helps us to determine both how much morality should demand of us and also how blameworthy we are when we fail to meet those demands.

1. Two tasks of moral theory

Begin with two intuitive pictures of morality, describing two kinds of tasks moral theory might have. First, on one view, a central task of moral theory is to provide the ultimate standard for us to live up to. If you hold this view, you might also believe some or all of the following things about moral theory:

Moral theory should evaluate our actions. A good moral theory should tell us what’s praiseworthy, blameworthy, and morally neutral. A good moral theory should tell us what the best life is, morally speaking. We should be able to learn from moral theory who the moral exemplars are. We may have no hope of

approximating the good moral character of a Jesus or a Buddha, but moral theory identifies why these are exemplars of virtue. It does not condescend to us by pretending that perfect virtue is possible for us; identifying perfect virtue will often show us just how far we have to go. Moral theory shows us how far others have to go, as well; it gives us ways to evaluate our own and others' behavior. It abstracts away from our flaws and biases, since bowing to these limitations cheapens morality. While moral theory provides us with moral reasons to do or not do certain things, those reasons are not necessarily reasons that are accessible to us. Because we are clouded by our own biases and lack of understanding, we may not ever be able to appreciate all the moral reasons, much less act on them. But seeing how distant morality is may inspire us to make progress. We can see that we have a lot to do to become better people, and we may look to the moral exemplars as examples of what we should try to become. If morality were closer to us, we might become complacent about our shortcomings. When moral theory sets the ultimate standard, it is morality 'in the sky;' it need not take note of the facts on the ground.

But on a second view, a different picture of morality emerges. Here, a central task of moral theory is to guide our actions. This view of morality might be linked to some or all of the following beliefs:

For moral theory to guide our actions, it must be able to tell us what to do given the kinds of people we are. If morality has nothing to say to flawed and imperfect people, then we can't make practical use of it. This means that morality should in some ways be constrained by our weaknesses and limitations. When we look at the actions that are, given our flaws, genuinely open to us, moral theory should be able to tell us whether one is better than another. It should be able to govern our interactions with other flawed people. This is moral theory that we can use in our deliberations: we can understand the moral reasons of this kind of moral theory, and we have the potential to act on them. This kind of moral theory is livable; we can follow at least some of its prescriptions, because it gives us advice that's suitable for us as we are. It can inspire us to be better, because it can give us the hope of acting decently. A morality that is further away may make being a good person seem hopelessly remote. Morality 'in the sky' can't guide our actions if it's too remote from what we are actually like. For this, we need morality 'on the ground.'

Perhaps you aren't persuaded that one or the other of these is really a task of morality. Many people aren't. But the thing is, they disagree. Some people I've discussed this with think that moral theory clearly has the first, and only the first, task: it should set us an ultimate standard, tell us the truth about morality. Anything else is 'advice column' morality. Other people think that moral theory clearly has

the second, and only the second, task: it should tell us what to do. If moral theory can't guide our actions, morality really isn't its subject at all. This disagreement is itself, I think, instructive. We may not universally accord both of these tasks of moral theory primary importance, or even any importance at all. Yet both of these views of what moral theory should do are picking up on something important and, to at least some people, intuitive about the nature of morality.

And yet these two tasks often seem to be mutually incompatible. In some cases, perhaps we are perfectly capable of living up to the standards set by the first view of moral theory. But many times we are not. Some quirk of our psychology or defect of our motivation gets in the way. When this happens, moral theory that only sets the ultimate standard cannot also give us guidance, beyond telling us to live up to a standard that's hopelessly remote. Moral theory that can give guidance is moral theory we actually have a chance of putting to use, but if it is our only moral theory, it is apt to seem complacent or parochial. It doesn't give us the full, often painful, truth about what we ought to be doing. It doesn't always push us to become better; sometimes it's too limited by the view from where we are right now. So these two tasks of moral theory often seem to be in direct conflict with each other. How can moral theory do everything we need it to do?

One way to explain the differences between these views of what moral theory is and what tasks it has is to look through the lens of the voluntarist constraint, the idea that 'ought' implies 'can.' Here's a preview of what I will argue: these views of moral theory are linked to different views about what 'can' means in each context—that is, whether it's tied to the very limits of what we can do or whether it hews more closely to our actual psychological and motivational makeup. Once we see the different kinds of 'can' at play in moral theory, we can begin to untangle this problem. But in order to do that, we must look more closely at the voluntarist constraint.

2. The voluntarist constraint

Most philosophers (although not all) accept the voluntarist constraint, that *ought implies can*. They believe, that is, that you can be obligated to do only those things that are possible for you to do. The voluntarist constraint is widely accepted partly because it strikes an intuitive chord. For one thing, moral theory that tells me to do things that are clearly impossible seems unfair. If my moral theory directs me to fly, then by failing to fly I am doing something wrong. But since I can't live up to this demand, morality is being unfair to me. For another, a moral theory that violates the voluntarist constraint is bad at giving us guidance. If my moral theory tells me

to fly, what guidance can that give me? Moral obligations that stretch far beyond what’s possible are too distant to give me information about what I should do.

Still, just because it’s plausible doesn’t mean it’s right. For every argument for the voluntarist constraint, someone has developed a counterargument.² I won’t adjudicate those arguments here, because I want to focus on a different question. If we accept for the sake of argument that ought implies can, then we need to know what that means. In particular, what does it mean to say we ‘can’ do something? One option is to hold that there is only one meaning of ‘can’ that is relevant to determining our obligations. But which one?

2.1 Possibility

When we say that someone can do something, we may mean any of several different things, including (but not limited to) the following:

2.1.1 Physical

I ‘physically can’ do something if it is possible for me to do it given facts about my physiology. Here, I mean solely facts about the body, as distinct from facts about the mind or brain. Under normal conditions, it is physically possible for me to attend a baseball game on a Sunday afternoon: I am strong enough and coordinated enough to get up off the couch, walk to my car, and walk from my parking spot into the baseball stadium. If, however, I were to become too weak to open the door to my house, or if someone locked me in, it would now be physically impossible for me to attend the game: I simply could not use my body in a way that would make it possible. This is the *thinnest* sense of possibility I discuss here: it relies on the fewest facts about us.

2.1.2 Psychological

A *thicker* concept is psychological possibility. I ‘psychologically can’ do something if it is physically possible *and* it is possible for me to do it given facts about my psychology—that is, given facts about my mind or brain. Various features of an individual’s psychology may render something psychologically impossible: certain pathologies, intelligence deficits, memory loss, and so on. Conversely, for people with unusually good minds, the psychological possibilities are less constrained: more things are psychologically possible for someone with a photographic memory than for me. Given the facts about my psychology, I can attend a baseball game on a Sunday afternoon: I don’t have a psychological block against going. If I had agoraphobia, however, it would be psychologically impossible. It would be impossible for me to make myself spend time out in an open space among the large

crowds of a baseball game. It would still be physically possible, of course. Nothing in the structure of my muscles, bones, joints, and so on prevents me from going to the game: if I had the right kind of mind, I could go. Someone else with an identical musculature but with a different brain could go to the game. But if I were an agoraphobe, I would not have the right kind of mind: given my condition, it would be *psychologically* impossible for me to go to the game. Our minds circumscribe the set of things we can do: the set of things that are psychologically possible is a subset of those that are physically possible.

2.1.3 Motivational

The final, and for our purposes, *thickest* modality is motivational possibility. I ‘motivationally can’ do something if it is psychologically possible *and*, somewhere in my motivational set, there is a first-order motivation to do that thing. By ‘motivations,’ I mean desires in a very broad sense—I can have the desire, or the motivation, to take out the trash or go to the dentist, even though no part of me looks forward to or enjoys those experiences. We do not act on every first-order motivation; one may be outweighed by another, stronger motivation. But for something to be motivationally possible, a first-order motivation must be present. I can want to want to take out the trash, but until that second-order desire activates a first-order desire, taking out the trash is motivationally impossible for me. When something is motivationally impossible, on the other hand, no first-order motivation exists.

In the case we’ve been discussing, given that I am not agoraphobic, it is *psychologically* possible for me to go to the baseball game. But, as it turns out, I just don’t enjoy baseball. I think it’s boring. In this case, I have no first-order motivation to go to the game. Absent some second-order motivation (all my friends are going; I have to meet another secret agent in a crowded location to hand off the evidence; my foreign friend wants to experience American sports), I cannot make myself go. There’s not a physical block; there’s not a block coming from my psychological abilities or inabilities. Going to a baseball game just for the love of baseball is *motivationally* impossible for me: I cannot want to do it.

In the same way that psychological possibility constrains physical possibility, motivational possibility constrains psychological possibility. Only some things are impossible for us because we lack first-order motivations. In cases of psychological impossibility, I might have strong first-order motivations yet still be unable to do something. An agoraphobe may dearly want to go to the game and yet be unable to, because some feature of her psychology is blocking her motivations from being effective.

2.2 Evaluating ‘can’

As the examples I’ve offered show, different modalities are relevant in different contexts. If I’m deciding how to spend a Saturday on my own, ‘I can go to the baseball game’ is false in the sense that matters. But if what I can physically do becomes relevant—if it matters whether you can reasonably request that I go to the game with you and your visiting British friend—then it matters that there’s a sense in which ‘I can go to the baseball game’ is true. ‘I can go’ and ‘I can’t go’ are both true at the same time, but they’re true on different senses of ‘can.’ So there is no one true analysis of ‘can’; there are many, and their truth varies by context.

Adding to the difficulty of understanding ‘can’ is that these modalities are not always clearly distinct. The lines *between* modalities are blurry. While my Midwestern relatives enjoy eating lutefisk, I gag whenever I try any; this reaction is probably mostly psychological, but it certainly seems to involve features of my physiology too.³ If my bad memory makes it psychologically impossible for me to remember your birthday, on the other hand, this psychological deficit might be connected to my motivations. These modalities are thus perhaps better thought of as regions on a continuum, without clear bright lines to distinguish them. At the same time, there may be finer gradations to make *within* modalities. What’s psychologically possible for me normally may be impossible for me when I’m under duress; there are finer gradations to make within each kind of possibility. But that the borders between these modalities are vague does not mean they are all of a piece. We may not be able to tell exactly where physical possibility bleeds into psychological possibility, but on only one of those senses can an agoraphobe truly go to the game.⁴

2.3 ‘Can’ and ‘ought implies can’

We might expect those who accept the voluntarist constraint to recognize the significance of the multiple meanings of ‘can.’ Not so.⁵ Instead, philosophers working on understanding the constraint have tended to pick one version of ‘can’ to tie obligations to. But interpreting the voluntarist constraint in this unitary way leads us into trouble, because a moral theory with only one voluntarist constraint cannot satisfactorily perform all of morality’s tasks. To see why, consider a recent disagreement between David Estlund and David Wiens about the uses of ‘ought implies can’ within political philosophy.

Estlund interprets the ‘can’ of ‘ought implies can’ in this way: ‘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’ (212). This is essentially physical possibility—on Estlund’s view,

we can even stay awake for four days straight (213).⁶ Estlund argues for this thin voluntarist constraint by appealing to the case of Bill the polluter (call him Selfish Bill), who claims that he is not required to refrain from dumping his garbage by the side of the road because he is too selfish to refrain (rather than because he has some phobia or compulsion) (220). Estlund claims that we still blame Selfish Bill for his garbage-dumping. Therefore, selfishness is not the kind of thing that can block a moral requirement; therefore, no motivational inabilities block requirements, even if they are inabilities all humans share (220). Because features of human nature do not block the requirements of justice, we wind up with a very thin sense of ‘can’ for the voluntarist constraint.

Wiens argues that the relevant sense of ‘can’ is much thicker: to him, ‘ought implies can will (in good faith)’ (339).⁷ He asks us to imagine Claudia, who makes repeated good-faith attempts to write a book but each time only writes a few pages before she gives up (Wiens 341). Can Claudia write a book? Wiens thinks she can’t; she’s proven that it is impossible for her to complete the sequence of events that results in a finished book. Similarly, whether Selfish Bill can avoid dumping his trash by the road depends on the reasons for his claimed inability. If Bill’s selfishness prevents him from making a good-faith effort to refrain from dumping, then he *cannot* refrain from dumping (and thus has no obligation not to dump); if he could, with the right good-faith effort, work past his selfishness, then he *can* refrain. While Estlund believes that moral theories ought to idealize away from human nature, Wiens says that we must take into account at least some of our inabilities when determining our obligations.

But neither of these single readings of ‘can’ tells the whole story. Wiens identifies an important distinction between good- and bad-faith motivational inabilities: there seems to be an important difference between someone who can’t do something because she is too selfish or cruel to do it and someone who can’t do something even when she tries in good faith to do it (345). This suggests that Estlund was too quick to reach the conclusion that motivational inabilities *never* block moral requirements; he may have traded unfairly on our lack of sympathy for someone who claims to be incurably selfish. Estlund’s view simply leaves us on the hook for too much—if no part of human nature counts against our being able to do something, then our obligations will far exceed what we could ever do. Wiens’s version of the constraint can give us guidance in a way that Estlund’s thinner, more remote constraint cannot. It is better at performing the second task of moral theory, that of giving us action guidance that we can use as we are.

But then again, Estlund is motivated by trying to figure out what justice is: ‘those to whom we owe justice do not lose their claim on us just because it might

turn out that we are not, perhaps even by our nature, disposed to deliver it’ (230). If Wiens’s Claudia has promised a publisher that she will write a book, then her inveterate procrastination is not enough to void her contract just because she’s tried in good faith to overcome it. Unlike if Claudia were physically unable to complete the book, there is a sense in which she still can complete the book, a sense which would make it possible if not for her flaws.⁸ Moral theory should let us hold others to account for their flaws even when they’ve made good-faith efforts to overcome those flaws. When Claudia has promised, we *should* say that Claudia ought (in some sense) to write the book; with Wiens’s ‘can,’ we cannot. Some opponents of the voluntarist constraint contend that it lets us off the hook for things we should be responsible for; Wiens’s version is especially susceptible to this criticism.⁹ Estlund’s version of the constraint can perform the first task, by providing an ultimate standard for us to live up to, regardless of our flaws; Wiens’s has a more difficult time doing so.

Thus both Estlund’s thinner and Wiens’s thicker specifications of ‘can’ seem to run into difficulties: neither can fully provide an ultimate standard and give us action guidance. What if we split the difference with a moderate voluntarist constraint, something like ‘ought implies psychologically can’? This only inherits, rather than solving, the problems with the extremes. When Claudia promises to finish the book but is psychologically unable to comply with this obligation, it seems entirely appropriate to hold her responsible. She has promised: and a promise doesn’t disappear just because your flaws make it impossible for you to carry it out. A moderate voluntarist constraint rules out holding that Claudia has an obligation (of any kind) in this case, since such an obligation would be psychologically impossible. This is too lax, inheriting the problems of a motivational voluntarist constraint.

At the same time, even a moderate voluntarist constraint can only provide limited action guidance. What should Selfish Bill do, given that it is motivationally impossible for him to refrain from dumping his trash? If moral theory only speaks to what’s psychologically possible for him, it ignores that his motivations make it impossible for him to do the right thing. Since thick voluntarist constraints are closely tied to your motivations, they can guide your actions in a wide variety of circumstances. Thin and moderate voluntarist constraints can’t do that. They may provide long-term goals for you—they may direct you to change your motivational structure—but they cannot give you immediate action guidance that draws on the motivations you have right now. This is too demanding, inheriting the problem of a physical voluntarist constraint. If we want morality that can both provide standards and give us guidance, we need an alternative strategy.

3. Moral theory with multiple voluntarist constraints

Rather than trying to find one single meaning of the voluntarist constraint that can account for everything we want morality to do, we must develop multiple versions of the voluntarist constraint using the different modalities discussed above. Just like multiple senses of ‘can’ are often true, we are often under multiple sets of obligations at once. Our moral theories (whether these are consequentialist, deontological, or virtue theories, or something else entirely) must be sensitive to multiple voluntarist constraints: ‘(this kind of) ought implies physically can,’ ‘(that kind of) ought implies psychologically can,’ and so on. This means that any moral theory will saddle us with multiple sets of obligations. We start with a thin voluntarist constraint, ‘ought implies physically can.’ This provides us with the ultimate standard, one which doesn’t yield to our psychological or motivational limitations. But in order to get more practical guidance, we look to moral theory that uses thicker voluntarist constraints. If we are thinking about how to improve our motivations so as to become morally better in the medium term, we will want to consider obligations drawn from what’s psychologically possible for us.¹⁰ When we need immediate moral guidance, we look at what’s motivationally possible for us and draw a third set of obligations from that set of possible actions.

What I’m offering here is more or less a contextualist analysis of ‘ought implies can.’ Contextualism is the metaethical view that the semantic content of certain ethical claims partially depends on some feature of the context in which those claims are made. Relevantly for our purposes, contextualists tend to endorse contextualist semantics not only for ‘ought’ but also for ‘can.’¹¹ This view is one way to explain how the two interact—that is, how to be a contextualist not just about ‘ought’ and ‘can’ but about ‘ought implies can.’ Because the semantic content of ‘can’ varies by context, and because we are granting the truth of the voluntarist constraint, the semantic content of ‘ought’ varies along with it. Thus we get different versions of the voluntarist constraint in different contexts. If the context is one in which purely physical possibility is conversationally salient, such as the context of ideal theory, then claims about the voluntarist constraint have a different content than if the context is one in which motivations matter as well.

Of course, there are different strains of contextualism. Some contextualists hold that the meaning of the moral ‘ought’ is relative to the speaker’s moral standards. This strain of contextualism is friendly to anti-realist metaethical views, although it does not strictly entail them.¹² But other strains are more friendly to realism, leaving open the possibility that at least sometimes the conversationally salient moral standards are mind-independent and objectively true.¹³ While I think my kind of contextualism is open to either realist or anti-realist metaethical views,

my sympathies lie with the realist ones. The meaning of ‘ought implies can,’ then, is determined by the relevant ability and by whatever moral theory turns out to be true. At any rate, my view does not imply anti-realist contextualism.

The contexts that are especially salient for our purposes—those related to motivational, psychological, and physical ability—range from less to more idealized. So this is one way that the distinction between ideal and non-ideal theory can solve problems in moral theory. In the ideal world, we would lack the limitations we have in the non-ideal world. We wouldn’t have the psychological or motivational inabilities that make it in some sense impossible for us to comply with morality’s demands. But in the actual world, these inabilities make it impossible for us to comply. That’s why we need distinct non-ideal theory that we can use to guide our actions in the non-ideal world. We derive our non-ideal obligations from our ideal obligations, but we modify them in light of our psychological and motivational limits. Because we have at least two sets of limits—psychological and motivational—we need different versions of non-ideal theory to carry out the different tasks of moral theory. So while there are two levels of theory—ideal theory, which provides the ultimate standard, and non-ideal theory, which provides action guidance—there are also multiple levels within non-ideal theory. Some non-ideal possibilities are further from the ideal than others.¹⁴

3.1 An example: act utilitarianism

Assume, just for now, that act utilitarianism is true.¹⁵ Then think about the duty of beneficence. On act utilitarianism, I have (at least) three sets of obligations of aid. First, I am obligated to render aid as far as is physically possible. While this level of possibility is often unhelpful for providing immediate practical guidance, it tells me what I should aspire to be able to do. Even the staunchest act utilitarian might admit that it is psychologically impossible to live up to act-utilitarian principles one hundred percent of the time—but as long as compliance with those principles is *physically* possible, I am nevertheless doing something wrong when I fail to live up to them. If aiding the needy consists in writing checks to Oxfam, then I am obligated to write checks to Oxfam until I reach the outer limit of what I am physically capable of doing.

On a non-ideal theory that includes facts about my psychology, I am obligated to render aid as far as my psychology will let me. Perhaps it is psychologically impossible for me to never favor my family over complete strangers. I may be able, as much as it hurts, to divide my money exactly equally, but I may still be unable to spend all of my time performing acts of beneficence, instead occasionally going to watch my kid play soccer. In this case, then, my

obligation is to help others as far as my psychology will let me but favor my own family where I must. What’s psychologically possible can often be useful as a medium-term goal: it isn’t as minimal as the set of obligations that are motivationally possible, but it also isn’t as unrealistic for us to consider as bare physical possibility is. That unrealistic obligation, tied to a physical version of the voluntarist constraint, still exists. It’s the source of my psychologically possible obligation, and it is in the background providing the ultimate standard of morality. But when it’s impossible in any sense, non-ideal theory steps in.

And finally, consider an even more non-ideal theory, which takes into account my motivation. For the purposes of immediate action guidance, this non-ideal theory tells us that I have the obligation to render aid as far as I can motivate myself to. This level, the most constrained set of obligations, gives immediate practical guidance. When I’m thinking about what to do right now, surely I should only consider whatever I can motivate myself to do. While other options exist, and may be in some sense obligatory, they are impossible in the sense of ‘can’ that matters for guiding actions. So if I’m deciding how to act in the short term, I should only consider those options that are motivationally possible for me. If I cannot even be motivated to give up buying lots of Christmas presents for my family, then I should aid as much as I can consistent with the amount of motivation I have to be beneficent. Ignoring my motivations will lead me into trouble—if I convince myself I can give away all my money, spend all of December researching various charities, and then find myself in the toy aisle on Christmas Eve, I will violate the obligations of both ideal theory (by not actually giving away my money) and non-ideal theory (by buying my children whatever lousy presents are still unsold).

Whatever the correct moral theory is, we start with the maximum obligations of that theory and then add thicker and thicker voluntarist constraints to get closer to what we are actually able to do. The different kinds of possibility thus act as successive filters. When we move from ideal to non-ideal theory, we filter out obligations that are impossible in a non-ideal context. These filters narrow down our options until, in order to get action guidance, we are only considering things we are currently motivationally able to do.

But this filter metaphor isn’t perfect. It might give the impression that since we are at every step removing obligations, surely we can never find ourselves with more obligations than we had at a thinner level of possibility. Yet as we move from ideal to non-ideal theory, we may find ourselves with more duties to make up for things we (in some sense) can’t do. If I have arachnophobia, perhaps I am not psychologically obligated to save a child from being bitten by a spider. But I may have (psychologically possible) duties to make up for my failure to save the child. I

might have duties to visit the child in the hospital, to pay some of his hospital bills, to undergo therapy to try to treat my phobia, and so on. And then there are our duties of self-improvement, which help to solve a potential problem with the view.

3.2 Transition to the ideal

That problem is this: it might seem like the theory that we actually use to guide our actions requires us to do hardly anything. If I am not motivated to donate money to the poor, or save a drowning child, or refrain from saying something mean to someone, then non-ideal theory tells me that I am not obligated to. Doesn't this let me off the hook too easily?

A first response is that even this very non-ideal theory is not as lax as it may seem. There is an important distinction between ‘can’t,’ ‘won’t,’ and ‘unlikely to’: ‘ought implies can’ has never meant ‘ought implies is easily within our grasp.’ We are off the hook for things we *cannot* be motivated to do, but we are not off the hook for things we can be motivated to do but *won’t* do or are *unlikely* to do. If I can bring myself to want to take out the trash, then I must do it, even if I would rather do something else. Laziness does not remove an obligation, even on the most non-ideal version of our moral theory. In order for non-ideal theory to be unable to command something, it must truly be motivationally impossible for us.

But even though motivational obligations can sometimes be demanding, sometimes they simply aren't. Maybe I truly cannot bring myself to want to give strangers equal consideration to my family members. In this case, a non-ideal moral theory which uses a motivational voluntarist constraint cannot require me to give equal consideration to strangers. This would be a problem if this were the only moral theory we had. But it's not. The two types of non-ideal theory show us why.¹⁶

The first type of non-ideal theory, *non-transitional* theory, tells us about our obligations in the non-ideal world. When I've been talking about non-ideal moral theory (about the parts of our moral theories that operate with psychological and motivational voluntarist constraints), I have mostly been talking about non-transitional theory. Even if I'm permanently stuck with my non-ideal motivations, there are some obligations I can act on right now.

The second type is *transitional* theory. This kind of theory tells us about our obligations to transition to the ideal. To see the difference, think about different reasons for punishing wrongdoers. We might punish for non-transitional non-ideal reasons: in the non-ideal world, people do the wrong thing, and we want to express society's disapproval. We might also punish for transitional reasons: we want people to learn from their mistakes in order to do better next time.

In the case of moral theory, I have transitional obligations to bring my motivations in line with what is psychologically and, ultimately, physically possible for me. We need these transitional obligations so that morality can carry out its first task, of providing us with the ultimate standard to live up to. Because there’s a sense in which I can complete any action that is physically possible for me, it remains a real obligation in that sense even if it is psychologically or motivationally impossible for me. These more demanding obligations don’t disappear just because of my motivational or psychological defects.

This helps to answer the worry that motivational moral theory is not demanding enough. I am obligated to try to become the kind of person who is motivationally able to comply with these other obligations. I should try to rearrange my motivational structure to comply with the obligations that are psychologically possible for me; I should try to improve my psychological structure in order to comply with the obligations that are physically possible for me. If I don’t, I am ignoring some of my moral obligations. Because moral theory with a motivational voluntarist constraint is not the entirety of our moral theory, we are not reliant exclusively on non-ideal theory.

In order to fulfill my obligations to transition from non-ideal moral theory to more ideal moral theory, I must use the stock of motivations I currently have in order to become a better person. In the case of fulfilling my obligations of beneficence, for example, it might be motivationally impossible for me to fulfill those obligations directly. But I may be able to trick myself into living up to those obligations. Perhaps I could get my spouse to set up a monthly transfer of money from my art fund to Oxfam. Over time, I might come to realize that I don’t miss that money; this realization might lead me to develop the motivation to carry out my charitable giving myself. If it’s motivationally impossible for me to be nice to someone I intensely dislike, maybe I can bargain myself into it by saying just one decent thing to that person and then leaving the room. As I get used to tricking myself into doing things I cannot be motivated to do, I can become habituated into adopting those motivations. As I work within the motivations I have, better motivations may become possible.

For some people, though, this transition may truly be impossible. Even if it is physically and psychologically possible to be a better person, it is completely motivationally impossible. Those people are off the hook—they are not obligated to become better people. For someone to be let off the hook in this way, it would have to be the case that he is completely devoid of all motivation to become a better person. More than that, that person would have to be completely devoid of all motivations that could *become* motivations to be a better person. It’s already

questionable whether someone who is completely devoid of even potential moral motivation can be held morally responsible. But most of us are capable of becoming better people, and so most of us are required to try.

So our transitional obligations are what require us to become better people and keep this framework for morality from being too lax. But the process is often not simple. Sometimes it will be easier to change certain physical attributes than certain motivational ones. I might be highly motivated to become a construction worker, in order to build homes for those who don't have them, but lack the minimum strength needed. In order to get the job in construction, I'll have to begin a weight-training regimen. Given that I'm highly motivated, I may find that the changes in my physical abilities come easily. On the other hand, as we have just seen, some people may never change their motivations if they see no reason to. That is, it may be easier to remove some filters at the top, at the physical level, than at the bottom, at the motivational level. At the same time, some physical and psychological limitations are irremediable no matter what our limitations are. But every change begins with changes to our motivations—in order for me to start lifting weights, I have to be motivated to do so. All removal of obstacles to the performance of our moral duties requires us to have the correct motivations in place, but some motivations are easier to correct than others.

Another complication is that our obligations at the various levels (ideal theory, transitional non-ideal theory, non-transitional non-ideal theory) may conflict. A case of conflict between my ideal and non-ideal obligations might look something like this: ideal theory requires me to refrain from blowing large sums of money at the blackjack table. But I am not motivated to quit gambling, so non-ideal theory does not obligate me to quit. On a particular Saturday night, I have promised a friend I'll go to a casino with him. If I stay home instead, I'll strengthen my motivation to quit ever so slightly, but I'll have broken a promise. Here, it seems that my ideal and non-ideal obligations are in conflict: ideally, I would never gamble, but since I am incapable of fulfilling that obligation, I am obligated to keep my promise.

So what should I do? Keep my promise or refrain from gambling? Remember that we've already said that it is motivationally impossible for me to act ideally. Because I cannot be motivated to quit gambling entirely, ideal theory does not provide immediate action guidance. So the only immediately possible obligations I can have with respect to this ultimate standard are obligations of transition. Given the choices open to me, this conflict between ideal and non-ideal obligations boils down to a conflict between transitional and non-transitional obligations. Should I keep my promise to my friend? Or should I refuse, upsetting

my friend, as part of a project of increasing my motivations to give up gambling later?

The details about how to resolve this kind of conflict will depend largely on the moral theory we accept. If utilitarianism is true, questions about transition will be answered by figuring out what will bring about more utility in the long term. Perhaps I am required to break my promise to my friend now so that I can quit gambling later. If deontology is true, transition to the ideal will properly be constrained by moral considerations such as basic rights. I am probably not allowed, all things considered, to murder someone now, even if that helps me fulfill some obligation of ideal theory later. We have an idea of the kinds of considerations that will provide all-things-considered guidance when obligations conflict: which moral theory I accept, what my options are, and how they help or impede my transition to ideal morality. But filling in the blanks will require us to say more about the content, not just the structure, of moral theory.

3.3 How demanding is the ideal?

One argument sometimes made against certain moral theories is that they are too demanding. What critics making this argument often mean is that the moral theory requires people to do things that are unrealistic to expect of people given what we are like. Flanagan makes an objection of this kind against act utilitarianism: act utilitarianism cannot be correct, he argues, because we can't do the moral math it requires. It is impossible for us to, with our limited time and cognitive resources, continually compute all possible actions we could take, their consequences, and the relative utility of each of these consequences (Flanagan 33-4). Another stock objection to act utilitarianism is that, in defiance of our psychology, it requires us to give all people exactly equal consideration. That's not something that most people are capable of doing.

But once we divide moral theory up into ideal and non-ideal versions, this kind of objection no longer gives us good grounds for dismissing a moral theory. When we are deciding which *ideal* theory of morality (ideal utilitarianism, deontology, virtue ethics, etc.) is correct, we cannot bring in facts about what is psychologically or motivationally impossible for *non-ideal* people to do. The ultimate standard of morality may not yield to our flaws; this would cheapen morality. If ideal morality demands more than we're psychologically or motivationally capable of, that's a flaw we should try to overcome, not a flaw in our ideal theory. So act utilitarianism cannot be dismissed on the grounds that Flanagan gives (33-34). Kantian deontology cannot be dismissed on the grounds that it is impossible for people to act solely from a motive of duty (Flanagan 36). If

people are physically capable of doing these things, then ideal morality can require them to do so. Questions about what our obligations are given our psychological and motivational inabilities are now questions for non-ideal theory. If a non-ideal theory doesn't correctly take our psychological or motivational limitations into account, that isn't a mark against its counterpart ideal theory; it's a sign that we should calibrate the non-ideal filters differently.

But this doesn't mean that the most extreme version of moral theory is necessarily the right one. Moral theory may make less extreme demands not because people are unable to comply with extreme demands but because morality *should not* make extreme demands even on those who *can* comply with them.¹⁷ Where Flanagan's objections to act utilitarianism don't work, Scheffler shows how a different kind of strategy can succeed. Scheffler argues that moral theory ought to take our personal integrity into account—it ought to give us some room to care about our own concerns and projects, even out of proportion to their impartial value (1982, 21). Because act utilitarianism requires us to maximize impartially good consequences, it can't do this. What we need, Scheffler thinks, is a view that includes a prerogative to give one's own interests this kind of extra value.

Notice that this kind of view need not rely on any claims about inability at all.¹⁸ Its claim is that even people who are psychologically and motivationally capable of becoming act utilitarians should not be morally required to do so. Scheffler's theory is less demanding than utilitarianism not because of claims about what we *can* do but because of claims about what a good life for us is like and what is properly important to us: morality should not alienate us even if its demands are possible (in all senses) for us. This ideal theory is different from ideal act utilitarianism, and subsequent non-ideal theories will be similarly distinct: we arrive at them by applying the thicker senses of possibility to each ideal theory. The disagreement between Scheffler and act utilitarians now centers on which *ideal* theory makes the right demands of *ideal* agents.

The differences between Scheffler and Flanagan help to illustrate which kinds of demandingness complaints can legitimately be made against ideal theories. It is not legitimate to reject ideal moral theories on the grounds that non-ideal agents cannot comply with their demands. We can certainly reject non-ideal moral theories on those grounds. If non-ideal theory with a motivational voluntarist constraint requires me to do something motivationally impossible for me, I should revise my theory so that it can provide me with usable guidance. But if my ideal theory requires me to do something motivationally impossible, I cannot reject it for that reason: I must instead try to become able to live up to its demands.

Still, in thinking about the kinds of demands ideal moral theory can make, it is legitimate for me to deliberate about what ideal morality should be like. What kinds of demands is it appropriate to make of ideal agents? What are the appropriate limits of morality even where compliance is not an issue? Ideal moral theories can come with more or less extensive demands. What they have in common is that they arrive at those demands not by looking at what people are capable of but at what morality may reasonably ask of people, regardless of their capabilities. Our arguments about what ideal moral theory should be like cannot be predicated on what non-ideal agents are capable of.

3.4 Blame

History is full of people who acted in ways which we regard as morally abhorrent but which seemed perfectly legitimate at the time. The men (and women) who worked against women’s suffrage were committing grave injustices, but many of them sincerely believed that women ought not be allowed to vote. These people seem to have done something wrong, and yet an opponent of women’s suffrage in the 19th century seems far less blameworthy than a 21st-century opponent would be.

When we divide up moral theory into ideal and non-ideal versions, we get different sets of obligations corresponding to different versions of the voluntarist constraint. This means that if someone fails to fulfill an obligation of ideal theory, blame may simultaneously be appropriate and inappropriate. It is appropriate to blame someone insofar as he could have fulfilled an obligation, but it is inappropriate to blame him insofar as he could not. People for whom the obligations of ideal theory are fully possible are fully blameworthy if they fail to support women’s right to vote. But for those for whom the obligations of ideal theory are impossible, we need a different story.

Now, perhaps some earlier opponents of women’s suffrage didn’t have the epistemic ability to become aware of the injustices they were committing; like physical and psychological ability, epistemic ability may affect the shape of our obligations.¹⁹ But with the advent of the women’s suffrage movement, epistemic inability seems like a less plausible explanation.

Setting aside, then, questions of epistemic ability, questions about psychological and motivational ability remain. Given the views of society at the time, it required an abnormal moral courage to accept gender equality; many of us are just not that psychologically strong. On non-ideal theory, we cannot blame the opponents of women’s suffrage, since standing up to societal norms exceeded their psychological ability. But ideal theory gives us room to blame them. They should

have been the kind of people who could be motivationally and psychologically able to treat women as equals. They should have had the strength of character to listen to the arguments of the suffragists and to accept them, even if that meant going against the grain of society. We can blame them for falling short in this way, even while acknowledging that their shortcomings were due to psychological and motivational inabilities.²⁰

This means that, although we can still blame these opponents of women’s suffrage, we must blame them less than those who commit comparable wrongdoing today. To borrow from Watson, 19th-century opponents of women’s suffrage are responsible (and thus blameworthy) mostly in the *attributability* sense (229). We can attribute these agents’ attitudes toward women’s suffrage to their values and beliefs. But it would be (mostly) wrong for us to hold these agents *accountable* for their beliefs, since it would have been motivationally or psychologically impossible for them to believe otherwise. There’s only a thin, physical sense on which these agents could have supported women’s suffrage. On the other hand, a 21st-century opponent of votes for women, someone for whom such a thing would be motivationally possible, is fully responsible in both the attributability and the accountability senses.

When we accept ideal and non-ideal versions of blame, then, we do not wind up with the kind of permissive relativism that would never blame people for doing things they wrongly believed to be right.²¹ We can still attribute to them actions that were wrong. We can even hold them accountable in a relatively thin sense, because there was a thin sense in which doing otherwise was possible for them. We also do not wind up with the excessively harsh condemnation that it would be unfair to give to people who were, by and large, not actually moral monsters (at least, not more so than the rest of us). If ideal obligations are too demanding, it can’t be because imperfect people fail to live up to them. But if we admit of multiple kinds of blame, to go with the multiple sets of obligations we have, then we can blame historical wrongdoers (and anyone else who fails to live up to the obligations of ideal theory) only in the ways that are appropriate given their circumstances.²²

4. Conclusion: how this resolves the two tensions

The early sections of this paper were devoted to laying out two tensions in moral theory: first, the tension between moral theory’s setting an ultimate standard and its ability to provide action guidance; and second, the tension between different interpretations of the voluntarist constraint. I then argued that we should adopt

ideal and non-ideal moral theory using multiple versions of the voluntarist constraint.

It should be fairly obvious how this would resolve the second tension, between different voluntarist constraints. We don't have to decide between voluntarist constraints: we can have, and in fact need, them all. If we are trying to figure out the ultimate standard of morality or justice, as Estlund is, then we should use a thin voluntarist constraint. If we are trying to figure out what we should actually do in a particular set of circumstances, one in which there are psychological or motivational constraints on what we can do, then we should make the voluntarist constraint thicker. In general, different types of possibility are relevant to different situations, and the same thing is true in moral theory. The tension between different voluntarist constraints disappears when we see that different constraints are relevant to different things we want moral theory to do.

That means that the resolution to the second tension shows us how to resolve the first. To determine the ultimate standard for judging actions, we look at ideal theory, which operates with a thin voluntarist constraint. But to guide our actions, we use non-ideal theory, which operates with thicker voluntarist constraints. The kind of moral theory that can guide our actions must take account of our motivational and psychological limitations. Whatever ideal moral theory turns out to be right—utilitarianism, deontology, or something else—it will need non-ideal counterparts. If we recognize that there are obligations that are motivationally and psychologically impossible, but physically possible, then we can see that there are more ideal versions of morality we should strive for.

Department of Philosophy
Rhode Island College

Works Cited

- Agule, C. (2016). ‘Resisting Tracing’s Siren Song.’ *Journal of Ethics and Social Philosophy* 10, pp. 1-24.
- Benson, P. (2001). ‘Culture and Responsibility: A Response to Moody-Adams,’ *Journal of Social Philosophy* 32, pp. 610-620.
- Bjornsson, G. and Finlay, S. (2010). ‘Metaethical Contextualism Defended,’ *Ethics* 121, pp. 7-36.
- Chrisman, M. (2015). *The Meaning of ‘Ought’: Beyond Descriptivism and Expressivism in Metaethics*. Oxford: Oxford University Press.
- Chuard, P. and Southwood, N. (2009). ‘Epistemic Norms without Voluntary Control,’ *Noûs* 43, pp. 599-632.
- Dowell, J.L. (2011). ‘A Flexible Contextualist Account of Epistemic Modals,’ *Philosophers’ Imprint* 11, pp. 1-25.
- Estlund, D. (2011). ‘Human Nature and the Limits (If Any) of Political Philosophy,’ *Philosophy and Public Affairs* 39, pp. 207-237.
- Finlay, S. (2014). *Confusion of Tongues*. Oxford: Oxford University Press.
- Fischer, J.M. and Tognazzini, N. (2011). ‘The Physiognomy of Responsibility,’ *Philosophy and Phenomenological Research* 82, pp. 381-417.
- Flanagan, O. (1991). *Varieties of Moral Personality*. Cambridge: Harvard University Press.
- Graham, P. (2011). ‘“Ought” and Ability,’ *Philosophical Review* 120, pp. 337-382.
- Howard-Snyder, F. ‘Ought Implies Can,’ in H. LaFollett (ed.) *International Encyclopedia of Ethics*. Hoboken: Blackwell Publishing, pp. 1-9.
- Jay, C. (2013). ‘Impossible Obligations Are Not Necessarily Deliberatively Pointless,’ *Proceedings of the Aristotelian Society* 113, pp. 381-388.
- Kekes, J. (1984). ‘“Ought Implies Can” and Two Kinds of Morality,’ *Philosophical Quarterly* 34, pp. 459-467.
- King, A. (2014). ‘Actions That We Ought, but Can’t,’ *Ratio* 27, pp. 316-327.

- Kratzer, A. (2012). ‘The Notional Category of Modality,’ in A. Kratzer, *Modals and Conditionals*. Oxford: Oxford University Press, pp. 27-69.
- Kratzer, A. (1977). ‘What “Must” and “Can” Must and Can Mean,’ *Linguistics and Philosophy* 1, pp. 337-355.
- Levy, N. (2003). ‘Cultural Membership and Moral Responsibility,’ *The Monist* 86, pp. 145-163.
- Mizrahi, M. (2009). “‘Ought’ Does Not Imply ‘Can,’” *Philosophical Frontiers* 4, pp. 19-35.
- Pittard, J. and Worsnip, A. (2017). ‘Metanormative Contextualism and Normative Uncertainty,’ *Mind* 126, pp. 155-193.
- Rawls, J. (1999). *A Theory of Justice*. Rev. ed. Cambridge: Belknap Press of Harvard University Press.
- Rosen, G. (2004). ‘Skepticism about Moral Responsibility,’ *Philosophical Perspectives* 18, pp. 295-313.
- Scheffler, S. (1992). *Human Morality*. New York: Oxford Press.
- Scheffler, S. (1982). *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford: Clarendon Press.
- Silk, A. (2017). ‘Metaethical Contextualism,’ in T. McPherson and D. Plunkett (eds.) *Routledge Handbook of Metaethics*. London: Routledge.
- Simmons, A. J. (2010). ‘Ideal and Non-Ideal Theory,’ *Philosophy and Public Affairs* 38, pp. 5-36.
- Sinnott-Armstrong, W. (1987). ‘Moral Dilemmas and “Ought and Ought Not,”’ *Canadian Journal of Philosophy* 17, pp. 127-139.
- Sinnott-Armstrong, W. (1984). “‘Ought’ Conversationally Implies ‘Can,’” *Philosophical Review* 93, pp. 249-261.
- Stern, R. (2004). ‘Does “Ought” Imply “Can”? And Did Kant Think It Does?’ *Utilitas* 16, pp. 42-61.

- Stocker, M. (1971). “Ought” and “Can,” *Australasian Journal of Philosophy* 49, pp. 3-16.
- Streumer, B. (2003). ‘Does “Ought” Conversationally Implicate “Can”?’ *European Journal of Philosophy* 11, pp. 219-228.
- Van Someren Greve, R. (2014). “Ought,” “Can,” and Fairness,’ *Ethical Theory and Moral Practice* 17, pp. 913-922.
- Vargas, M. (2005). ‘The Trouble with Tracing,’ *Midwest Studies in Philosophy* 29, pp. 269-91.
- Värynen, P. (2006). ‘Ethical Theories and Moral Guidance,’ *Utilitas* 18, pp. 291-309.
- Vranas, P. (2007). ‘I Ought, Therefore I Can,’ *Philosophical Studies* 136, pp. 167-216.
- Wedgwood, R. (2006). ‘The Meaning of “Ought,”’ in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*, vol. 1. Oxford: Clarendon Press, pp. 127-60.
- Wolf, S. (1988). ‘Sanity and the Metaphysics of Responsibility,’ in F.D. Schoeman (ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. New York: Cambridge University Press, pp. 46-62.
- Watson, G. (1996). ‘Two Faces of Responsibility,’ *Philosophical Topics* 24, pp. 227-248.
- Wiens, D. (2016). ‘Motivational Limits on the Demands of Justice,’ *European Journal of Political Theory* 15, pp. 333-352.

¹ Rawls came up with this distinction (or at least gave it its name), but there is now a flourishing literature on these two types of theory, especially as they relate to justice (see esp. 214-220).

² See Vranas and Howard-Snyder on arguments for the voluntarist constraint. For arguments against the constraint, see, among others, Mizrahi, Graham, and King. Sinnott-Armstrong (1987) and Stern accept weaker readings of the constraint but argue that the traditional formulation of the constraint is too strong (on this, see also Streumer’s response to Sinnott-Armstrong). For counterarguments specifically directed at the idea that ‘ought implies can’ makes morality fairer, see Van Someren Greve and Sinnott-Armstrong (1984).

³ Lutefisk is cod treated with lye until it becomes a kind of fishy Jell-O. Thanks to Steve Swartzler for this vivid example.

⁴ And there are other kinds of possibility that might be relevant too—nomological, metaphysical, or even logical. Whether you think these are relevant to moral theory will

depend on your precise view of the voluntarist constraint. I’ll focus on the three kinds of ability I discussed above, but I’ll flag places other kinds might be relevant.

⁵ Chuard and Southwood (614) mention in passing the existence of multiple kinds of ‘can,’ but they do not pursue this point. Kekes and Jay both note, but neither pursues, the point that limitations beyond physical inability might constrain obligations if the voluntarist constraint is true.

⁶ This sense of possibility appears to exclude from consideration both psychological and motivational features (Estlund 230-35).

⁷ More precisely: ‘A person is able to (can) [do something] if and only if she successfully completes a sequence of acts that manifests [that thing] in a sufficiently high proportion of the possible worlds at which she repeatedly makes a good-faith attempt to complete a sequence of acts that conduces to [that thing]’ (Wiens 347). Graham has a similar sense of ‘can’ in mind; on his view, an arachnophobe cannot touch a spider (342).

⁸ What if Claudia *made* herself physically incapable of completing her book (say, by locking herself in a room from which she couldn’t escape)? If it’s literally physically impossible, then the voluntarist constraint tells us that there can be no obligation. But we can still hold Claudia responsible for her past actions, especially those she might have undertaken with an eye to limiting her moral obligations in the future.

⁹ For a version of this contention, see Mizrahi. Stocker writes that the voluntarist constraint ‘would almost certainly be uninterestingly false if considered in light of *psychological* inability,’ arguing that the plausible sense of the constraint is a physical sense of ‘can’ (311).

¹⁰ Philosophers sometimes make a distinction between ‘deontic’ and ‘evaluative’ kinds of ‘ought’ (Howard-Snyder 1; Chuard and Southwood 601). While evaluative ‘ought’ statements (‘Life ought to be fair’) do not make demands on us, deontic ‘ought’ statements (‘I ought to tell the truth’) do. All versions of ‘ought’ on the non-ideal-to-ideal continuum make demands of some kind or another, so the ideal/non-ideal distinction is a distinction *within* the category of deontic ‘ought.’

¹¹ For ‘can,’ see Kratzer (1977); for ‘ought,’ some places to look are Chrisman, Finlay, Bjornsson and Finlay, Silk, Wedgwood, and Dowell.

¹² Examples can be found in Bjornsson and Finlay, Finlay, and Silk, among others.

¹³ Here see Wedgwood, Dowell, and Pittard and Worsnip.

¹⁴ I mentioned above that other kinds of ability (nomological, metaphysical, logical) may be relevant in some contexts. If they’re relevant here, then we might get obligations at an even more idealized level than physical possibility—obligations that we have because they’re metaphysically possible even though they’re physically impossible. Obligations based on these senses of possibility probably violate many understandings of the voluntarist constraint. But we should leave open here whether there are relevant kinds of ability beyond the one I’m picking out as the most ideal, and so even more ideal theories.

¹⁵ This framework applies to any moral theory, but I’m using act utilitarianism here because of its simple structure.

¹⁶ For more on the transition from non-ideal to ideal, see Simmons.

¹⁷ See Estlund (222-23).

¹⁸ It’s true that Scheffler sometimes makes suggestions about what is psychologically possible for us (see 1992, 68). But he need not *rely* on those suggestions, since space for personal prerogatives can come out of what morality ought to ask of us, not out of what we *can* do.

¹⁹ See Värynen for a discussion of epistemic possibility (‘certain kinds of information pertaining to the act are available’)(302).

²⁰ This does not imply a *tracing* view of moral responsibility, since claims about responsibility are not traced back to some previous time. Any time there’s a sense in which a wrongdoer can fulfill a moral obligation, he’s blameworthy if he doesn’t; any time he can’t, he’s not. We don’t need to know how that inability or inability came about. This avoids the complications of tracing views (see Vargas and Agule).

²¹ For examples of this kind of permissive relativism, see Benson, Levy, Rosen, and Wolf.

²² Because there are multiple levels of ideal and non-ideal theory, we will potentially confront many different attributability/accountability combinations in assigning blame. For support for this from a Watson-style view, see Fischer and Tognazzini.