

2016

## Moral Failure — Response to Critics

Lisa Tessman  
*Binghamton University*, ltessman@binghamton.edu

---

### Recommended Citation

Tessman, Lisa. 2016. "Moral Failure — Response to Critics." *Feminist Philosophy Quarterly* 2, (1). Article 6. doi:10.5206/fpq/2016.1.6.

## ***Moral Failure—Response to Critics***

Lisa Tessman

### **Abstract**

I briefly introduce *Moral Failure* as a book that brings together philosophical and empirical work in moral psychology to examine moral requirements that are non-negotiable and that contravene the principle that “ought implies can.” I respond to Rivera by arguing that the process of construction that imbues normative requirements with authority need not systematize or eliminate conflicts between normative requirements. My response to Schwartzman clarifies what is problematic about nonideal theorizing that limits itself to offering action-guidance. In response to Kittay, I defend my rejection of “ought implies can,” and consider the implications of the concept of unfair moral requirements.

**Keywords:** moral failure, nonideal theory, moral psychology, constructivism, moral demandingness

In *Moral Failure: On the Impossible Demands of Morality* (2015), I argue that there can be situations of unavoidable moral failure, or, put differently, that there can be impossible moral requirements, moral requirements that contravene the principle that “ought implies can.” We apprehend impossible moral requirements intuitively, when we grasp both that “I must” and “I can’t.” The most obvious cases of unavoidable moral failure are a result of moral dilemmas, in which either option that one can choose involves wrongdoing. Both deontologists and consequentialists rule out moral dilemmas, but I argue that there are genuine dilemmas, by drawing a distinction between negotiable and non-negotiable moral requirements; only those that are negotiable, and which concern values that can be substituted or compensated, are canceled when overridden in the resolution of the conflict. Non-negotiable moral requirements remain binding even when they become impossible to fulfill. Other cases of impossible moral requirements arise in situations where it is a person’s need or vulnerability that gives rise to a moral requirement; when the need is basic and significant, but also impossible to meet—for instance, when someone needs for an irreplaceable loss to be replaced or for an irreparable harm to be repaired—the moral requirement to which it gives rise is also impossible to fulfill.

I draw on empirical work that supports a dual-process model of moral cognition; we have both an automatic, unconscious, affect-laden intuitive system and a consciously controlled reasoning system, and while we can make our judgments through either process, the empirical work shows that most of our moral judgments are made intuitively, and that reasoning tends to take place after the judgment has been made, to justify the judgment to others (Haidt 2001). I show that while the reasoning processes referred to by deontologists and consequentialists cannot yield judgments that we are impossibly required, we can indeed make intuitive judgments that we face an impossible moral requirement, because “ought implies can” cannot insert itself into the intuitive process the way it can in the reasoning process.

One of the main concepts in the book is the idea of the *unthinkable*. To make a judgment that an action is morally *wrong* is just to judge ‘I must not do it.’ To make an intuitive judgment that an action is *unthinkable* is to judge both that ‘I must not do it’ and that ‘I must not even consider doing it’ (by ‘consider’ I mean engage in justificatory reasoning about). That is, it matters *how* one makes the decision, and it matters because it matters to *us*, namely to people who want certain affect-laden, automatic, intuitive responses from each other. The idea of the unthinkable is behind Bernard Williams’s famous example of the rescuer who has “one thought too many” when he considers whether morality permits him to prioritize his wife or requires that he decide impartially between rescuing his wife or a stranger first (1981, 18). The wife “might have hoped” that it was an affect-laden intuitive judgment that prompted her husband to rescue her first (18). Harry Frankfurt captures the same idea in his characterization of love as involving a volitional necessity about actions that protect or care for one’s beloved; love for him constitutively involves finding certain actions—such as actions that threaten one’s beloved—to be unthinkable (Frankfurt 1982, 2004).

In such cases—in other words, cases where what we value in another person’s response to us is not just their action but their emotions as well, emotions that motivate them to act in a certain way toward us without questioning, calculating, or justifying—we (morally) must rely on our intuitive judgments. There is further empirical work on “sacred values” showing that individuals and moral communities tend to treat a set of values as sacred, as infinitely valuable, so that any consideration of sacrificing the value (even if one chooses not to sacrifice it) is forbidden. That is, in the case of sacred values it is wrong to think one thought too many—to rationally consider rejecting the value (Tetlock et al. 2000; Tetlock 2003).

Because our judgments regarding impossible moral requirements are intuitive judgments, and these requirements may be requirements not to do the unthinkable, they must be shielded in a certain way from reasoning that involves comparing or weighing values—that is, reasoning that involves considering doing

what should be unthinkable. But this means that our usual ways of double checking on our intuitions—such as reflective equilibrium (Rawls 1971) or even Margaret Urban Walker’s transparency testing (Walker 1998, 2003)—might actually violate the values about which we must not think one thought too many. In some way we need to leave these values unquestioned, but leaving them unquestioned is also very dangerous, because our intuitions can be ideologically shaped (that’s what transparency testing is meant to address) and because social groups sometimes sacralize values in such a way that they view other values or certain other people as evil threats to sacred values, threats that they then believe must be eradicated (Graham and Haidt, 2012).

My book, more broadly, is an account of what morality is and how we construct it—I’ll say more about this in response to Lisa Rivera’s comments. It also considers what more we might want from morality than action-guidance, namely that we might want to assess actions as constituting moral failures, even when we cannot avoid performing these actions—I’ll say more about this in response to Lisa Schwartzman’s comments. And it enters discussions of what is known as the question of moral demandingness, arguing that if other people’s needs and vulnerabilities can be sources of moral requirements for us, and sources that in some cases we respond to with the intuitive sense of “I must,” then morality, far from being moderate the way that many theorists want to argue that it is, can be impossibly demanding—I’ll say more about this in response to Eva Kittay’s comments.

I would like to thank all three critics for their wonderful, thought-provoking comments. Each of them, through their own work, has contributed over the years to my thinking, and I greatly appreciate hearing their reflections on my work.

## Response to Critics

### Response to Lisa Rivera

Lisa Rivera offers what I find to be an illuminating interpretation of *Moral Failure*, characterizing it as implicitly identifying, and addressing, two different conflicts: the “theory vs. phenomena” conflict and the “explanation vs. prescription” conflict. I do believe, as Rivera suggests, that both conflicts can be avoided or resolved. The theory vs. phenomena conflict exists for theories that try to systemize—but because I don’t think a theory must (or should) thoroughly systemize, at least not in such a way that it reduces plural and conflicting values into a harmonious unity, one can avoid the tension between theory and phenomena by abandoning the quest for a unified system (sorry, Parfit, but there is more than one mountain to climb). As Rivera points out, giving up on systemizing may leave us

facing conflicting moral requirements, and if these moral requirements are non-negotiable, we will also face unavoidable failure. If we acknowledge this, and accept a moral theory that has conceptual space for impossible moral requirements and unavoidable failure, then there need not be any conflict between (such a) theory and moral phenomena.

The explanation vs. prescription conflict is more complicated. This conflict would clearly arise for moral realists if they were to accept the sorts of scientific explanations that I do, because then it would be terribly mysterious why our normative experiences, as outcomes of the processes that can be scientifically explained, would happen to match the moral prescriptions that follow from the independently existing moral facts that realists posit. Sharon Street demonstrates this in her brilliant debunking of moral realists' accounts of morality, in "A Darwinian Dilemma for Realist Theories of Value" (2006). But I'm not a moral realist, so the question is whether there is still any explanation vs. prescription conflict for an anti-realist of a constructivist sort.

I don't think that there necessarily is; whether or not this conflict exists will depend on what the constructivist takes the process of construction to be, and whether that process is one that will tend to produce authoritative moral prescriptions that cannot be reconciled with the subjective experiences (the phenomena)—such as experiences of requirement—that we (or scientists) might describe or explain. My position is that we prescribe in part because we have certain experiences, and engage in certain evaluative processes, though a description or explanation of these experiences or processes does not entail any particular prescription. We (or neuroscientists) might explain what neurological processes take place in the brain when we prescribe. We (or evolutionary anthropologists/biologists/ psychologists) might explain what evolutionary processes led to our being the kinds of creatures who will tend to experience certain things normatively or prescriptively, or what psychological mechanisms we have evolved to motivate our prescriptive behavior. Since we can study the processes involved in prescribing, we can then explain or describe them, but it is not the explanation that does any of the work of prescribing or of lending authority to our prescriptions—we do that when we prescribe or when we place our confidence in certain prescriptions. There is no reason why explaining how and why we prescribe would preclude prescribing, or even take any of the punch out of our prescriptions. (We might also note that there is likely to be an explanatory gap: the subjective experience of requirement, or the "what it is like" to have a normative experience or to prescribe, cannot be completely captured by a scientific explanation. But the fact that we cannot completely capture this in our explanations does not diminish the force of our prescriptions).

If we express our prescriptions in the form of normative theories, and if these theories try to systemize beyond what the phenomena explained (partially) by the scientific theories can support, then we are back to the phenomena vs. theory tension. But our prescriptions, whether or not they take the form of normative theories, need not systemize, and need not eliminate conflicts (though if they do not, our prescriptions may be impossible to carry out).

We could include, under the umbrella of prescribing, a wide range of evaluative acts, any of which may contribute to the construction of morality. I believe (unlike most other constructivists) that the construction of morality is not limited to being a justificatory endeavor, since some moral requirements come to carry authority without being justified. Prescriptions are produced in some sense from both first-order judgments and second-order judgments, and these might differ from each other in how much authority we take them to carry. Thus we could include as a prescriptive act the act of forming an immediate, intuitive judgment that “I *must*” or (to put it in imperative form) “Do this!” —accompanied by a feeling of necessity or requirement about what one is to do. We could also include the act of producing a verdict as a result of a conscious reasoning process, such as “Push the large person off the footbridge.” First-order judgments, though, are not necessarily authoritative, since in many cases we will end up without confidence in a first-order judgment and this lack of confidence will deprive the prescription of its moral authority. This is why we might want to distinguish between prescriptions that are backed with moral authority and prescriptions that are not, but we must keep in mind that there is not just a single route to imbuing a prescription with authority. In the case of sacred values, the tendency will be to automatically endorse and have confidence in one’s first-order intuitive judgments; and, I have argued, this is the only way in which sacred values may (permissibly) be judged. The first-order judgment that “I must protect this sacred value” and the second-order judgment that I must not even consider rejecting the first-order judgment (lest I have one thought too many) come together automatically. In other cases, rational reflection, or the social process as characterized by Haidt’s social intuitionist model (Haidt 2001), may indeed be an important part of the process of coming to have confidence in a prescription and granting it moral authority.

Rivera writes: “What I am responding to in situations—what I believe makes me required—seems to have no immediate explanatory or causal connection to the metaethical construction/social justificatory process. It is hard to see how the process plays a role in the moment I believe myself required.” I would modify this somewhat; the process of construction is wider than the “social justificatory process” and so the two should not be conflated. The process of construction may include justification—typically performed socially—but it is not limited to

justification. The “moment I believe myself required” —that is, the intuitive grasping of “I *must*” —is also part of the process of the construction of morality.

In any case, our prescriptions, and the moral authority that they carry, are not somehow derived from our explanations; one cannot, after all, derive an *ought* from an *is*. Instead, the prescriptions come from *us*, when we do any of the things that count as prescribing; we *construct* them through the wide range of processes that are involved in the construction of morality. My account of the construction of morality portrays these processes as a good deal messier than they appear to be in other constructivist accounts, because I don’t think that the processes are always, or should always be, reasoning processes, and I don’t think that they yield a consistent set of prescriptions.

Some constructivist accounts may create a tension or conflict between explanation and prescription, if the scientific explanations reveal a plurality of relevant experiences, all of these experiences are taken as the raw materials for the process of construction, and the process is supposed to yield a consistent set of prescriptions. Here there will be tension because the plurality will not be easily reduced into a consistent set, so the explanation will point to plurality and the prescription will not “fit” this plurality. As soon as one relinquishes the hope for a single correct prescription, the tension dissolves.

What remains is a morality that is thoroughly contingent. It is contingent on what we value and imbue with authority, and thus on what sorts of processes and structures lead to our having confidence in the authority of any particular moral requirement. These may include different cognitive processes through which we have normative experiences, and different social arrangements that structure how we interact to create shared understandings of what is required of us. Morality must, then, be relative to these contingencies.

### **Response to Lisa Schwartzman**

Lisa Schwartzman is right that I am interpreting “nonideal theory” to mean theory that is action-guiding in the actual, nonideal world, and that in doing so I am following some other critics of ideal theory, as well as being guided by Rawls’s own understanding of the distinction between ideal and nonideal theory. However, I want to clarify that I do not deny that nonideal theorists may do other things as well—for instance, Schwartzman’s own work in *Challenging Liberalism* (2006) works out a methodology for a kind of nonidealizing epistemology, aimed at understanding oppression as structural (while ideal theory, in being ideological, obscures this fact). The reason that I focus on nonideal theory’s centering of action-guidance is because this is where I think a problem arises.

It is not that I object to offering appropriate action-guidance, since we clearly need this kind of action-guidance, and especially action-guidance that has taken into

consideration what the actual wrongs are that need to be addressed through our actions (rather than action-guidance that is inappropriate—and potentially ideological—because it is based on an idealized picture of the conditions in which the action is to take place). Rather, what I object to is the thought that, in offering action-guidance, we answer the question, “What is the right thing to do?” In other words, I object to what I take to be the false sense that if we make the right action-guiding decision—which just means that we choose the best of the available alternatives, all of which might be awful—then we have not failed.<sup>1</sup> I want action-guidance *and* action-assessment (Stocker 1990), so that we recognize the awfulness of what we do even when we do the best that we can, and I would like for this recognition to replace the righteousness we might otherwise have about what we do.

Schwartzman asks why I don’t challenge Rawls’s own characterizations of the division between ideal and nonideal. Rawls understands nonideal theory to concern situations of noncompliance or partial compliance with principles of justice. I don’t think that noncompliance or partial compliance *must* be interpreted, as Schwartzman interprets it, as referring to the actions of *individuals* (“individual bad behavior”), though it is true that there is a lot of literature that does interpret it this way, such as literature on the moral demands faced by “compliers” in light of others’ noncompliance (e.g. Murphy 2000). I understand noncompliance with justice to include any form that injustice takes, and because injustice tends to take structural forms, I take nonideal theory, even when understood as theory about situations of noncompliance or partial compliance with principles of justice, to indeed concern itself with “fundamentally oppressive structures of social hierarchy and power.”

Schwartzman also wonders why I might be interpreting nonideal theory as action-guiding theory. Those whom I follow in doing this include Amartya Sen (2006, 2009) and Charles Mills, though primarily the position that Mills expresses in *Contract and Domination* (Pateman and Mills 2007), rather than in “‘Ideal Theory’ as Ideology” (2004). Mills’s critique of ideal theory is emphasized in “‘Ideal Theory’ as Ideology” while his development of what nonideal theorists should do is the focus of his chapters in *Contract and Domination*. Mills’s claim in “‘Ideal Theory’ as Ideology” is that the absence of descriptive accounts of actual conditions, and the presence of idealizations, serve an ideological purpose. With this I absolutely agree (and so I do not problematize it in *Moral Failure*). What it seems to suggest that nonideal theorists should do is to correct for the mistakes of ideal theory by foregrounding descriptive accounts of the actual and by scrutinizing normative ideals to determine whether they are truly worthy. In *Contract and Domination*, Mills is explicit in his claim that we ought to utilize nonideal theory to arrive at “prescriptions for remedial

---

<sup>1</sup> This idea is also expressed by Martha Nussbaum (2000).



justice in a racist social order” (Pateman and Mills 2007, 114)—that is, action-guidance in the nonideal, actual world. He proposes that we do this by imagining a hypothetical “later position” (in place of the “original position”), which puts parties in a position to choose principles of reparative justice, to choose the actions that should take place to compensate for a history of (racial) injustice. In other words, Mills insists here that the descriptive account, combined with a vision of the normative ideal for which to strive, enables us to offer appropriate action-guidance.

Sen, too, is quite explicit about nonideal theory being action-guiding; in particular, it should guide us in how to make comparative action-guiding choices, where the comparison is between different options that are available in an actual, nonideal society. Sen (2006) contrasts what he calls “transcendental” and “comparative” theories, and advocates comparative theories. He claims that transcendental theories—theories that identify what a perfectly just society would be like—are neither necessary nor sufficient for making comparative choices between available alternatives, all of which are nonideal. Without commenting directly on his argument for these claims, I want to point out one way in which I differ from Sen: I *do* think that we need what Schwartzman calls “normative ideals” (Sen’s “transcendent ideals”), *but for a different purpose than for guiding action*.

This brings into view the mistake that I think nonideal theorists tend to make when they offer action-guidance that they believe aims at their normative ideals. The problem is not with their offering action-guidance, nor is the problem with their holding normative ideals. The problem is the tendency to not question the relationship between the two. As the “general theory of second best” (Lipsey and Lancaster 1956, discussed in Goodin 1995) demonstrates, the best *possible* action (call this the second-best, since an impossible action is the first-best) is not necessarily the action that one would choose if one were trying to reach the normative ideal, the first-best. Robert Goodin illustrates this point by asking us to suppose that the ideal is to be served gelato with chocolate sauce. If I am only served fettuccine and not gelato, then instead of coming as close as possible to the ideal by having fettuccine with chocolate sauce, I will choose as the second-best (that is, the best that is possible), fettuccine with marinara sauce. To give an example related to politics rather than fine dining: if we have only a racist society (fettuccine), we would be better off choosing racially based affirmative action (marinara) than choosing colorblindness (chocolate sauce). But let’s not lose sight of the fact that the first-best, the normative ideal (the gelato with chocolate sauce), would be a racially egalitarian society with no further need for racial reparations and so one where colorblindness would be appropriate. The normative ideal here cannot be action-guiding, but it should be recognized as that which is the truly *worthy* (i.e. first-best) ideal. The best action-guiding decision is to pursue racially based

affirmative action over colorblindness, but only because we know we cannot reach the normative ideal.

We must not lose sight of the gap between the (non-action-guiding) first-best—which I call a “worthy ideal”—and the (action-guiding) second-best—which I call a “feasible ideal” (2015, 199). Nonideal theory is very good at telling us what to do in the nonideal conditions we are in, but this is usually accompanied by the illusion that if we follow the prescription, we will be getting closer to our (worthy) normative ideals, as well as the illusion that we won’t be doing anything wrong. However, we *won’t* necessarily be getting closer to the ideal; sometimes we are getting farther away from it, and I count this as a failure, even when it is the best thing we can do.

Thus I do stand by my claim that there are two problems with nonideal theory:

- 1) It inadvertently idealizes the moral agent by portraying the agent as someone who always can do the right thing simply by making the right action-guiding decision, and so as someone who does not face unavoidable moral failure.
- 2) It suggests that what nonideal theory tells us to do is *good* (rather than *not good enough* but better than the alternatives). For instance, every time one chooses (whether one does this as an activist, as a member of a progressive institution, as an applied ethicist, and so on) how to allocate resources, one makes trade-offs, often in situations where making trade-offs is unacceptable because the sacrificed value is one that cannot be substituted or compensated; in other words, one violates a non-negotiable moral requirement. To think of this as doing the right thing is to inappropriately adapt (borrowing language here from Serene Khader’s term, “inappropriately adaptive preferences” [Khader 2011]) one’s normative expectations downward, labeling something as good even if it is something terrible.

### **Response to Eva Kittay**

The most basic question that Kittay poses to me is: why am I so intent on rejecting “ought implies can”? Why should “ought implies can” be tossed out as one of the dogmas of moral theory? After all, there seem to be so many good reasons—several of which Kittay points out—to affirm that “ought implies can.” For one thing, it keeps us sane and wards off some of the pain of seeing ourselves as having failed. Kittay tells us, in fact, that the very idea that she could be required to do something impossible seemed crazy to her from the moment her therapist pointed it out. I don’t think it is crazy to judge oneself to be impossibly required. I don’t think it’s *rational*, but neither do I think that makes it crazy. I don’t think it’s rational in the sense that I don’t think it’s a *reasoned* judgment. But I believe that we should have

confidence in the authority of some of our automatic, intuitive moral judgments—including judgments about impossible moral requirements—even when they must sit alongside other judgments that are inconsistent with them. If we don't keep, and endorse, some of our intuitive judgments of being impossibly required, we lose some morally valuable ways of responding to each other. In some cases, when we rule out impossible moral requirements we rule out a certain kind of love.

But nevertheless Kittay's point—that there are many good reasons to affirm that “ought implies can”—is well taken. That's why I think that the process of construction of morality cannot produce a consistent set of judgments. I think it can produce *both* the judgment of impossible requirement and the judgment that “ought implies can.” Both are valuable. Because of this, I do think that “ought implies can” functions as a dogma of moral theory, which prevents us from apprehending impossible moral requirements and recognizing the value of doing so.

Kittay wonders why I don't join Bernard Williams (1973) in rejecting the agglomeration principle instead of rejecting “ought implies can.” I doubt I would have argued for rejecting “ought implies can” if my own intuitive judgments did not favor rejecting it, so the next question is what I should do with these intuitive judgments—should I dismiss them, the way I would want, for instance, to dismiss intuitive judgments that exhibit implicit bias? I find myself extremely resistant to giving up my intuitive judgments about impossible moral requirements. My resistance is based on my sense that in some cases the only alternative to the act that I judge to be impossibly required is an act that is unthinkable.

If I make an intuitive judgment that something is impossibly required, I may also make an automatic second-order judgment that the first-order judgment about an impossible moral requirement must not be questioned through justificatory reasoning. But it is only through such reasoning that the principle that “ought implies can” could be applied, and that inferences could be made from it. Justificatory reasoning, in which I consider rejecting the requirement, would constitute “one thought too many” if the alternative to the required act is unthinkable.

The clearest cases, for me, of experiencing some option as unthinkable, take place in the context of a certain kind of love, which Harry Frankfurt presents beautifully. Someone who experiences this kind of love “finds that he [*sic*] *must* act as he does” (Frankfurt 1982, 264). As Frankfurt puts it, “An encounter with necessity of this sort characteristically affects a person less by impelling him into a certain course of action than by somehow making it apparent to him that every apparent alternative to that course is unthinkable” (263). It is not that there is any requirement or duty *to* love, or to love in this way, or to love anyone in particular; rather, *if* you love in this way, then your love itself gives you a reason, and leads you to find some actions (such as actions that protect the beloved) to be absolutely

required and other actions (such as actions that sacrifice the beloved) to be unthinkable. So why must we sometimes make our judgments about what we are required to do intuitively rather than through reasoning? Because love requires it; because *if* we love someone in a certain way, then some things will be unthinkable. We'd be unable to even bring ourselves to consider doing them, to have that one (reasoned) thought too many.

Unthinkability serves as a sort of impossibility; an unthinkable act is impossible *in the sense of* being unthinkable. This means that the alternative to an unthinkable act must be absolutely required (though it may also be impossible). This is why love—and the unthinkability of certain acts, which we grasp when we love in a certain way—can necessitate that we acknowledge ourselves to be bound by impossible moral requirements. If choosing not to protect someone whom I love is unthinkable—and impossible in the sense of being unthinkable—then, even if I cannot in fact protect them, I still must do so. To not find myself to be impossibly required would reveal that I take it to be thinkable not to protect, which in turn would reveal an absence of a certain kind of love.

If we want to live in a world in which one can love and be loved in the way that comes with the automatic, decisive sense that we *must* do certain things—a sense that is paired with the realization that all alternatives to what we must do are unthinkable—then we must accept a world in which we don't necessarily reject our own judgments that we're required to do the impossible. That is why I am so resistant to giving up my judgments that I am impossibly required. I don't want to be the wife of that rescuer who thought the unthinkable when he thought one thought too many (nor do I want to be that rescuer).

I argue in the book that it is in more than just relationships of love that we value people's finding some acts to be unthinkable. In a variety of situations that go well beyond particular relationships of love, part of what we want from other people *is* their emotional response or their affect-laden intuitions; we want them to have an immediate feeling of horror, for instance, at the thought of pushing the large person off the footbridge in the Push version of the trolley problem (Tessman 2015, 66–67). As Philip Tetlock's empirical work finds, people do tend to want and expect other people to make a certain class of moral judgments intuitively, and to find certain actions to be unthinkable—that is, protected from a process of rational consideration (Tetlock et al. 2000; Tetlock 2003).

I want to return for a moment to why Kittay may find it to be simply crazy to speak of requiring the impossible. It *would* be crazy to judge someone who has failed to do the impossible to be morally blameworthy in the same way that someone would be if they had failed to do something that they *could* do. Unavoidable moral failure is not the same as avoidable moral failure. Blame, including self-blame or guilt, is not exactly right in the case of unavoidable moral

failure. When I have failed to do what is impossibly required, I feel a pained recognition of the failure, but this is not the same as guilt.

While Kittay is right that I treat the principle that “ought implies can” as a dogma to be questioned, Kittay’s portrayal of me as treating the action-guidingness of moral theory as dogma is not exactly accurate. I believe that one of the things that moral theory must do is offer action-guidance. What I find to be a dogma of moral theory is the notion that moral theory begins and ends with action-guidance; under the influence of this dogma, moral theorists conflate “what is to be done” with “what it is right to do.” What I and others, such as Michael Stocker (1990) and Christopher Gowans (1994), have emphasized against this dogma is that even after the action-guidance and the action have taken place, there is more for the moral theorist to say. For instance, following Stocker, I believe that we can still *assess* the action that we have decided is the action to be done: we can still judge that performing this action constitutes a failure. What is the point of doing this? The point, here, is to match the subjective experience of unavoidable failure. When I find myself facing what I intuitively judge to be an impossible requirement, and I find myself failing to fulfill it, it is alienating to turn to moral theories that have no conceptual space to account for this phenomenon. Offering the concept of impossible moral requirements, and enabling us to claim that we have unavoidably failed, are non-action-guiding functions of moral theory. Moral theories that perform these non-action-guiding functions are, I believe, better than those that merely guide action, even though I do agree that action-guidance must be, and usually is, the central task of moral theory.

Lastly, I turn to Kittay’s discussion of how I treat her vulnerability model of moral requirements, which she develops in *Love’s Labor* (1999). According to this model, if someone is vulnerable to or dependent on a moral agent, and the moral agent is positioned to be able to respond to the dependent’s needs, then the moral agent may be morally required to do so. This is true even if the moral agent did not voluntarily choose to be in such a position. However, if the moral agent was *coerced* into being in such a position, then according to Kittay everything changes. I interpret Kittay to be saying that a vulnerability-responsive moral requirement simply does not arise in cases where the moral agent has been coerced into being in a position where someone is dependent on her. In her comments, Kittay challenges my claim that, to the contrary, one *can*—due to bad (systemic) moral luck—be saddled with a vulnerability-responsive moral requirement, regardless of whether it is due to coercion that one is positioned to be required to respond to someone’s need or vulnerability. This claim is important to my view in part because situations in which one has been coerced into a position whereby one is required to care for a dependent might be dilemmatic (depending on whether there are conflicting requirements, and on whether both requirements are non-negotiable).

I think that an important distinction—brought to my mind by Anja Karnein’s argument in “Putting Fairness in Its Place: Why There Is a Duty to Take Up the Slack” (2014)—will illuminate our disagreement. The distinction is between cases in which the agent who is doing the coercing is also the dependent, and cases in which the agent who is doing the coercing is a different person than the dependent. Karnein’s point is that in the latter kinds of cases, there can be *unfair* moral requirements; the “place” of the unfairness (or other wrongdoing) is in the relationship between the coercer and the coerced moral agent, and it does not affect the fact that a moral requirement arises out of the need or vulnerability of the dependent, as long as the dependent is a third party.

Karnein’s point is made in the context of debates about moral demandingness and is targeted at theorists, such as Liam Murphy (2000), who argue that the demands of morality are limited, and thus moderate, because one cannot be required to do more than one’s “fair share.” These theorists, whom Karnein dubs “Fair Shares Objectors,” insist that in situations where some parties (such as the global poor) are in need and other parties (such as those who are well-off) are in a position to meet that need, the moral requirement to meet the need must be distributed fairly amongst those who can meet the need. According to the objection, in nonideal conditions in which some of those who are able to meet the need refuse to do so—call them the noncompliers—those who are willing to comply with the requirement still cannot be required to do *more* than their fair share, that is, the amount they would have to do under ideal conditions of full compliance. In other words, they cannot be required to pick up the slack.

Karnein disagrees, charging that the objection “fails to distinguish between two separate issues: first, what individual agents owe to the third party, and second, how fellow duty bearers ought to relate to each other” (Karnein 2014, 594); furthermore, “while the problem of fairness arises chiefly in the latter context, individual agents’ duties to third parties are determined independently of how fellow duty bearers ought to relate to each other... there will be cases in which what individual agents owe to third parties amounts to taking up the slack of noncompliers, even though this is very unfair” (594–595). The key to Karnein’s argument is separating two distinct relationships: the relationships of fellow duty bearers to each other (where the unfairness takes place when some are noncompliers) and the relationship of any one duty bearer to a third party. In the relationship between the duty bearer and the third party, if the third party is dependent, that dependency gives rise to a moral requirement to meet the need. Putting fairness in its place means putting it in the relationship between fellow duty bearers.

Karnein illustrates her point by imagining different scenarios in which some people need rescuing (such as children in danger of drowning), and there are several

potential rescuers, some of whom are noncompliers in the sense that they do not comply with the requirement to do their share of the rescuing. The details vary in order to capture some subtleties, but the basic idea is this: If two children are drowning and you, the only other competent swimmer around, walk away, you have treated me unfairly (burdening me with your share of the rescue work), and you have wronged the child whom you should have rescued (because you “recklessly endangered” her [Karnein 2014, 604]); nevertheless, I am now faced with two drowning children who have not wronged me in any way, and thus I must rescue both of them. Your behavior becomes part of the background of bad luck that has put me in the position of being required to rescue two children.

In order to translate Karnein’s argument to Kittay’s vulnerability model, let’s conceive of the parties as follows. Suppose that Agent A and Agent B are the potential duty-bearers (or in Kittay’s language, agents who potentially have vulnerability-responsive obligations), and C is a dependent. For instance, we can imagine that A and B are C’s parents, and that there is no one other than A and B who can or will care for C. Karnein’s point will be that unfairness between A and B—let’s suppose A is unfair to B by not doing his share of the work of caring for C—does not change the fact that if C still needs care and A is not doing it, the requirement falls on B. A’s noncompliance functions like luck for the purpose of the relationship between B and C; it is B’s bad luck to face C’s need—and the requirement to meet it—alone.

We can even imagine A to be more horrible than someone who is simply unfair. Let’s suppose B never wanted to have a child but was manipulated by A into doing so; let’s suppose A is abusive to B, and isolates B in the home, cutting her off from support, and undermining her chances of earning a living that would enable her to leave A; let’s suppose A spends his money on alcohol and drugs, or on luxuries for himself, providing B with so little money that B can only meet C’s basic needs by sacrificing some of her own. Clearly, A is wronging B in some very serious ways that go beyond simply being unfair; to use Kittay’s language (and to drastically understate things), A is not treating B like some mother’s child. Still, Karnein’s point holds. A wrongs B (and A also wrongs C). But this wrongdoing plays the role of bad moral luck—as Karnein puts it, it functions just like a force of nature (605)—in the way that it affects the obligations that B has to C (again, assuming that there is no one else who could care for C). Karnein is also careful to point out that A’s wrongdoing does *not* function like a force of nature in the way that it affects the relationship between A and B; there, it functions as the action of a responsible moral agent, A, who must be treated as an agent rather than as a force of nature, and thus can and should be held accountable (605–606). A owes all kinds of reparations to B. A also owes all kinds of reparations to C, because A has not met his vulnerability-responsive moral obligations to C. But none of this nullifies the

obligations that B has to C, which include the obligation to pick up A's slack. C, after all, is not the one who has wronged B; nothing about C's need has changed, and nothing in the relationship between C and B would inhibit that need from giving rise to a vulnerability-responsive moral requirement.

I'll make two points about this case. The first is a clarification to forestall misinterpretation. I am *not* saying that the correct action-guiding decision for B is necessarily to meet C's needs. Rather, if meeting her own needs—or, for instance, protecting her dignity—conflicts with meeting C's needs, then the best action-guiding decision for B may very well be to prioritize her own needs or rights. But if she does so, that decision will involve unavoidable moral failure, namely her failure to fulfill her vulnerability-responsive moral requirements toward C. The fact that protecting herself or her dignity conflicts with meeting C's needs does not serve to cancel the moral requirement to meet C's needs. If we follow Goodin (1985) in conceiving of B as having duties toward herself, then we can describe the situation as a genuine moral dilemma.

The second point is that it is crucial to the case (as for Karnein's cases) that the wrongdoer, A, is not the same person as the dependent, C. This may be what is behind Kittay's and my disagreement: Kittay has in mind two-person cases in which the dependent is the wrongdoer (such as the master, Johnny, in *Jubilee*), and I have in mind three-person cases in which the dependent is an innocent third party (such as the dying Auschwitz prisoners in Primo Levi's narrative). I think that in my reading of *Love's Labor* I did not pay attention to the fact that the cases that Kittay presents as cases in which coercion prevents a moral requirement from arising are primarily two-person cases in which the dependent is the wrongdoer. In cases like this, I am willing to agree with Kittay that the dependent/wrongdoer has no claim on the care of the person whom he wrongs. When we put fairness (or any other wrongdoing) "in its place" in such a case, its proper place (and the place of any wrongdoing and accountability for the wrongdoing) is in the very relationship in which a vulnerability-responsive moral requirement would otherwise arise. Part of the wrongdoer's being accountable for the wrongdoing may mean forfeiting the care of the person whom he has wronged. (Although I also agree with Kittay that even in such cases, it is possible that the relationship between the two parties *qua* human beings may permit a moral requirement to arise, while in the relationship between the two parties *qua* master and slave, a relationship that lacks "moral warrant," no moral requirement can arise).

But if the dependent is a third party rather than a wrongdoer, then this dependent has done nothing that would entail forfeiture of care; the innocent third party's need still triggers a vulnerability-responsive moral requirement. Thus the situation described by Primo Levi must be treated differently than the situation



described in *Jubilee*; the dying prisoners whom Levi feels impossibly required to respond to are a third party and have not wronged him in any way.<sup>2</sup>

When we put fairness in its place—and recognize that the source of unfairness or other wrongdoing is often distinct from the source of our moral requirements, and so the unfairness or other wrongdoing does not serve to prevent the moral requirements from arising—we can see that Karnein has given us a novel concept: the concept of *unfair duties*. There are moral requirements that it is unfair for us to be bound by, but that nevertheless are binding on us. Karnein does not intend for this concept to be applied in cases where meeting the requirement would be impossible—in fact, she explicitly rules this out, presumably because she accepts that “ought implies can” (see Karnein 2014, 594n4). However, I find that acknowledging that there can be unfair moral requirements makes impossible moral requirements more palatable, for I take it that part of the reaction against impossible moral requirements is due to their unfairness. After all, what could be more unfair than being required to do the impossible? It is terribly unfair to be set up for failure, and impossible moral requirements do exactly that. To this I say: yes, impossible moral requirements are totally unfair—but that does not diminish their authority.

## References

- Frankfurt, Harry. 1982. “The Importance of What We Care About.” *Synthese* 53 (2): 257–272.
- . 2004. *The Reasons of Love*. Princeton, NJ: Princeton University Press.
- Goodin, Robert. 1985. *Protecting the Vulnerable*. Chicago: University of Chicago Press.

---

<sup>2</sup> Kittay mentions the case of a woman who is pregnant by rape and wonders whether she has a duty to give birth. My response is, of course not. This is not because the dependent fetus is a wrongdoer, but rather because the fetus is not yet a person whose needs generate vulnerability-responsive moral requirements. If the woman does give birth, then she will face a requirement to find someone—not necessarily herself—to care for the child (obviously the rapist also has an obligation toward the child, but presumably he will be a noncomplier and will fail to fulfill this obligation, leaving the mother to pick up the slack). But the mere presence of the mother’s obligation toward the child does not entail that the right action-guiding decision for her is a decision to meet the obligation, because the obligation to find care for the child might be overridden by another requirement; in that case, however, the overridden obligation, which is non-negotiable, would not be canceled and so would leave a moral remainder.

- . 1995. "Political Ideals and Political Practice." *British Journal of Political Science* 25 (1): 37–56.
- Gowans, Christopher. 1994. *Innocence Lost: An Examination of Inescapable Moral Wrongdoing*. New York: Oxford University Press.
- Graham, Jesse, and Jonathan Haidt. 2012. "Sacred Values and Evil Adversaries: A Moral Foundations Approach." In *The Social Psychology of Morality: Exploring the Causes of Good and Evil*, edited by Mario Mikulincer and Phillip R. Shaver. Washington, DC, American Psychological Association: 17.
- Haidt, Jonathan. 2001. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108 (4): 814–834.
- Karnein, Anja. 2014. "Putting Fairness in Its Place: Why There Is a Duty to Take Up the Slack." *Journal of Philosophy* CXI (11): 593–607.
- Khader, Serene J. 2011. *Adaptive Preferences and Women's Empowerment*. New York: Oxford University Press.
- Kittay, Eva Feder. 1999. *Love's Labor*. New York: Routledge.
- Lipsey, R. G., and Kelvin Lancaster. 1956. "The General Theory of Second Best." *Review of Economic Studies* 24 (1): 11–32.
- Mills, Charles. 2004. "'Ideal Theory' as Ideology." In *Moral Psychology*, edited by Peggy DesAutels and Margaret Urban Walker, 163–181. Lanham, MD: Rowman and Littlefield.
- Murphy, Liam. 2000. *Moral Demands in Nonideal Theory*. Oxford: Oxford University Press.
- Nussbaum, Martha. 2000. "The Costs of Tragedy: Some Moral Limits of Cost-Benefit Analysis." *Journal of Legal Studies* 29 (2): 1005–1036.
- Parfit, Derek. 2011. *On What Matters*. Oxford: Oxford University Press.
- Pateman, Carole, and Charles Mills. 2007. *Contract and Domination*. Cambridge: Polity Press.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Schwartzman, Lisa. 2006. *Challenging Liberalism: Feminism as Political Critique*. University Park: Pennsylvania State University Press.
- Sen, Amartya. 2006. "What Do We Want from a Theory of Justice?" *Journal of Philosophy* 103 (5): 215–28.
- . 2009. *The Idea of Justice*. Cambridge, MA: Harvard University Press.
- Stocker, Michael. 1990. *Plural and Conflicting Values*. Oxford: Oxford University Press.
- Street, Sharon. 2006. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127: 109–166.
- Tetlock, Philip. 2003. "Thinking the Unthinkable: Values and Taboo Cognitions," *Trends in Cognitive Science* 7 (7): 320–324.

- Tetlock, Philip, Orië Kristel, S. Beth Elson, Melanie Green, and Jennifer Lerner. 2000. "The Psychology of the Unthinkable: Taboo Trade-Offs, Forbidden Base Rates, and Heretical Counterfactuals." *Journal of Personality and Social Psychology* 78 (5): 853–870.
- Walker, Margaret Urban. 1998. *Moral Understandings: A Feminist Study in Ethics*. New York: Routledge.
- . 2003. *Moral Contexts*. Lanham, MD: Rowman and Littlefield.
- Williams, Bernard. 1973. "Ethical Consistency." In *Problems of the Self*. Cambridge: Cambridge University Press.
- . 1981. "Persons, Character and Morality." In *Moral Luck*. Cambridge: Cambridge University Press.

LISA TESSMAN is Professor of Philosophy at Binghamton University. She teaches and does research in ethics, moral psychology, feminist philosophy, and related areas. Her work focuses on understanding how real human beings construct morality and experience moral demands, especially under difficult conditions. She is the author of *Burdened Virtues: Virtue Ethics for Liberatory Struggles* (OUP, 2005), and *Moral Failure: On the Impossible Demands of Morality* (OUP, 2015). She is currently completing a book entitled *When Doing the Right Thing Is Impossible*, which will be her first book written for a general audience; it is under contract with Oxford University Press.