

JOURNAL *of* ETHICS  
& SOCIAL PHILOSOPHY

VOLUME XX · NUMBER 2

*August 2021*

ARTICLES

- 119 Differentiating Disobedients  
*Chong-Ming Lim*
- 144 Scope Restrictions, National Partiality, and War  
*Jeremy Davis*
- 168 Defending the Epistemic Condition on Moral  
Responsibility  
*Martin Montminy*
- 188 Kant and the Problem of Unequal Enforcement  
of Law  
*Daniel Koltonski*

DISCUSSION

- 211 On Emad Atiq's Inclusive Anti-positivism  
*Kara Woodbury-Smith*

The *Journal of Ethics and Social Philosophy* (ISSN 1559-3061) is a peer-reviewed online journal in moral, social, political, and legal philosophy. The journal is founded on the principle of publisher-funded open access. There are no publication fees for authors, and public access to articles is free of charge and is available to all readers under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license. Funding for the journal has been made possible through the generous commitment of the Gould School of Law and the Dornsife College of Letters, Arts, and Sciences at the University of Southern California.

The *Journal of Ethics and Social Philosophy* aspires to be the leading venue for the best new work in the fields that it covers, and it is governed by a correspondingly high editorial standard. The journal welcomes submissions of articles in any of these and related fields of research. The journal is interested in work in the history of ethics that bears directly on topics of contemporary interest, but does not consider articles of purely historical interest. It is the view of the associate editors that the journal's high standard does not preclude publishing work that is critical in nature, provided that it is constructive, well-argued, current, and of sufficiently general interest.

*Editor*

Mark Schroeder

*Associate Editors*

Dale Dorsey

James Dreier

Julia Driver

Anca Gheaus

Errol Lord

Tristram McPherson

Colleen Murphy

Hille Paakkunainen

*Discussion Notes Editor*

Kimberley Brownlee

*Editorial Board*

Elizabeth Anderson	Philip Pettit
David Brink	Gerald Postema
John Broome	Joseph Raz
Joshua Cohen	Henry Richardson
Jonathan Dancy	Thomas M. Scanlon
John Finnis	Tamar Schapiro
John Gardner	David Schmidtz
Leslie Green	Russ Shafer-Landau
Karen Jones	Tommie Shelby
Frances Kamm	Sarah Stroud
Will Kymlicka	Valerie Tiberius
Matthew Liao	Peter Vallentyne
Kasper Lippert-Rasmussen	Gary Watson
Elinor Mason	Kit Wellman
Stephen Perry	Susan Wolf

*Managing Editor*

Jesse Wilson

*Copyeditor*

Susan Wampler

*Typesetting*

Matthew Silverstein



## DIFFERENTIATING DISOBEDIENTS

Chong-Ming Lim

ACTIVISTS WHO BREAK THE LAW on the basis of their conscientious—sincere and serious—moral or political convictions (henceforth “conscientious disobedients,” or simply “disobedients”) often (if not always) face the demand to differentiate themselves from “ordinary” criminals whose actions also violate laws but are not undergirded by conscientious convictions.<sup>1</sup> In general terms, this demand for disobedients to differentiate is not implausible—it serves an important function. Individuals who satisfy it are regarded as having a better (though nonetheless defeasible) claim to both the rights-based protections that are granted for conscientious action and to any putative legal excuse that may exist for conscientious breaches of the law. In practical terms, we have (defeasible) reasons to be—and indeed often are—more forgiving in our responses to and treatment of those whose violations of the law are undergirded by conscientious convictions, compared to those whose violations are not.<sup>2</sup>

1 The category of “conscientious disobedients” partially overlaps with that of “principled disobedients.” The former is differentiated on the basis of disobedients’ *convictions* (whether they are conscientious); the latter on disobedients’ *actions* (whether they are concordant with moral or political principles). An individual can be conscientious without being principled, and vice versa. Both categories may include those who behave civilly, uncivilly, directly, or indirectly, among others. For further discussions, see Brownlee, *Conscience and Conviction*, 18–27; and Delmas, *A Duty to Resist*, 21–46. For a discussion of how the requirement that disobedients be principled excludes or denigrates a certain class of resisters (especially the “lower class”), see Scott, *Weapons of the Weak*, 286–303. I set aside the class of individuals whose lawbreaking activity is motivated by basic needs.

The differentiation demand as I construe it differs from the requirement that political activists *never engage in any lawbreaking activity*. The latter is an implausibly narrow formulation of the differentiation demand and runs counter to the commonly held judgement (in most Western liberal societies) that activists can disobey the law in at least some circumstances without thus being no different from criminals. For further discussions of the differentiation demand, and how accusations of criminality are often used to discredit activists, see Lovell, *Crimes of Dissent*, 3–10; and Terwindt, *When Protest Becomes Crime*. For a discussion of how the figure of the criminal (especially as a racialized figure) has come to represent the most menacing enemy, see Davis, “Race and Criminalization.”

2 Brownlee, *Conscience and Conviction*, 7. I do not discuss the grounds for or practical implica-

In order to satisfy the differentiation demand, it is insufficient for disobedients to simply assert that their disobedience is undergirded by their conscientious convictions. To the extent that disobedients can make such assertions, so can criminals. Even if the disobedients' assertions are true, they do not secure the differentiation in the minds of other people. The differentiation demand concerns how disobedients present themselves and are perceived. It is not about whether their unseen mental states and motivations distinguish them (in some "objective" sense) from others who break the law. That is, the differentiation has to be secured from the perspective of their audience.<sup>3</sup> Observers—who do not have unmediated access to the "internal" states of disobedients—have to look for indications of disobedients' conscientious convictions. To secure differentiation, disobedients have to show that their appeal to those convictions is not just talk. One way of doing so is to behave in ways that are visibly distinct from those of the criminal and that indicate to others that they indeed conscientiously hold the relevant convictions.<sup>4</sup>

In some cases, the burdens involved in satisfying the differentiation demand may be onerous. These burdens have to be contextualized. Even if citizens and common institutions are prepared to accommodate or tolerate lawbreaking acts that are undergirded by conscientious convictions, they nonetheless have interests in avoiding being strung along by criminals who may falsely assert their possession of such convictions in their attempts to avoid punishment. Here, the thought is that in bearing these burdens, a disobedient shows herself to indeed have conscientious convictions.<sup>5</sup> Living in accordance with those convictions is so important to her that she is prepared to bear those burdens, and moreover may regard doing so as being on the whole worthwhile. And in bearing those burdens, she makes the conscientiousness of her convictions and actions plain for others to see—and thus differentiates herself from criminals who do not, and who are not prepared to, bear those burdens.

Within both public and philosophical discourse, the differentiation demand

---

tions of our partiality toward conscientious disobedients in this essay.

- 3 See Brownlee, "Reply to Critics," 727. Here, I take the audience to be the "general public," broadly construed. This is a simplification—the audience is not a monolith. Depending on which subset of the general public we are concerned with, the differentiation of disobedients from criminals may be either facile or nearly impossible. Some members of the general public may also make the differentiation demand on disobedients in bad faith, in their attempt to preserve the status quo. I set aside these complications for future work.
- 4 There are other ways of specifying the differentiation demand, without reference to individuals' conscientious convictions. See, for instance, Hannah Arendt's discussion of disobedience, which centers on the public and collective nature of civilly disobedient action (*Crises of the Republic*, 49–102).
- 5 This may be so even if she bears the burdens for purely strategic reasons.

is often understood as setting constraints on the actions that disobedients can engage in. For instance, and most commonly, a disobedient is regarded as falling afoul of this demand when she engages in “radical” actions such as arson, rioting, vandalism, or vigilantism, among others. Such actions are often regarded as failing to communicate the disobedients’ conscientious convictions, or even as being incompatible with such convictions. They are regarded not as conscientious disobedience but as mere criminal activity. For instance, the actions of the participants in the 2011 England riots and 2015 Baltimore riots—both protests against alleged police brutality against people of color—were denounced as simply criminal activity.<sup>6</sup> Similar criticisms have also been made of the 2020 Black Lives Matter protests and riots across the United States. Less radical or destructive actions are often also included within the category of radical actions—for instance, harassing political figures at their residences rather than workplaces, denying political figures service at businesses on the basis of their actions, or even engaging in covert and anonymous cyberattacks.<sup>7</sup> At the extreme, there are also those who judge any act of disobedience as indistinguishable from, or even worse than, ordinary criminality.<sup>8</sup> While their specifics vary, these criticisms are unified—they urge us to judge and treat those individuals as criminals who are undeserving of the protections or excuses typically afforded to conscientious disobedients.

In this essay, I argue that in some circumstances the differentiation demand can be satisfied by disobedients who engage in what are typically regarded as radical actions. In practical terms, this means that even disobedients who engage in actions such as arson, rioting, vandalism, or vigilantism can also successfully differentiate themselves from criminals.<sup>9</sup> The category of conscientious disobedients is potentially more inclusive than has been commonly assumed within public and philosophical discourse. Insofar as we think that conscientious disobedients should be judged and treated differently from criminals, we have reason to judge and treat disobedients who engage in these radical acts of disobedience differently from how we currently do.<sup>10</sup>

6 “House of Commons Hansard Debates for 11 Aug 2011,” Publications and Records, UK Parliament, accessed February 19, 2020, <https://publications.parliament.uk/pa/cm201011/cmhansrd/cm110811/debtext/110811-0001.htm>; Swaine, Jacobs, and Lewis, “Baltimore Protests Turn into Riots as Mayor Declares State of Emergency.

7 Beinart, “Left Wing Protests Are Crossing the Line”; Cochrane, “Sarah Huckabee Sanders Was Asked to Leave Restaurant over White House Work”; Thompson, “Hacktivism.”

8 These are views held by a number of judges. See Brownlee, *Conscience and Conviction*, 6, 156–58. As I suggest in note 1, this view is implausible.

9 For a controversial defense of looting as an instrument of political resistance, see Osterweil, *In Defense of Looting*.

10 Provoking disproportionate state response is often part of activists’ strategy—to lay bare

My argument proceeds as follows. In section 1, I briefly present the core features of Kimberley Brownlee's prominent account of the communicative principle of conscientiousness, which is one of the few systematic specifications and elaborations of the conditions constituting the differentiation demand. Over the next two sections—and partly in response to Brownlee—I articulate and defend two core aspects of my account, which provides a qualified defense of conscientious disobedients. In section 2, I argue that the communicative conditions should be characterized as *paradigmatically* true of those who *show* that they act on the basis of conscientious convictions, rather than as *necessary and sufficient* conditions for *having* those convictions. In section 3, I argue that while the conditions serve as practical tests of disobedients' convictions, disobedients' singular or even occasional failure of these tests need not threaten or eliminate their differentiation from criminals. They may still satisfy the differentiation demand if we adopt a holistic assessment of their persons and actions. In section 4, I consider the objection that my account imposes overly stringent constraints on disobedients. I conclude in section 5.

Before proceeding, two quick clarifications of the scope of my discussions are important. First, I focus on the conscientiousness—sincerity and seriousness—of disobedients' convictions and set aside the issue of their content.<sup>11</sup> The issues are distinct; an individual may conscientiously hold an abhorrent conviction. Of course, we may very well decide that those who conscientiously hold abhorrent convictions are no better (or perhaps even worse) than criminals. In which case, we may see the differentiation demand as applying only to those who do not hold such convictions. I take no stance on this issue here. I note only that in determining or judging that someone who conscientiously holds an abhorrent conviction is no different from a criminal, we would still need to consider how and whether they (or their actions) have securely indicated that they sincerely and seriously hold those convictions.

Second, I focus narrowly on the differentiation demand. This is distinct from the issue of whether disobedients behave justifiably or permissibly. A disobedient can be adequately differentiated from criminals yet behave impermissibly, or she may fail to satisfy the differentiation demand yet behave permissibly. While there is an extensive literature on the permissibility of disobedience, compara-

---

the violence in the system. My rehabilitation of disobedients does not rule out these strategies, and may even bolster them. The state's disproportionate response to lawbreakers who are recognized as disobedients may be even more frowned upon (and better galvanize action) than if those lawbreakers were regarded as criminals.

11 For a distinction between conscientiousness thus described and a morally nonneutral idea of conscience, see Brownlee, *Conscience and Conviction*, 7, 16–17.



tively much less attention has been directed at the issue of whether by engaging in such actions disobedients provide adequate indication of their conscientious convictions. Indeed, many recent discussions of disobedience—especially those that extend the discussions of defensive ethics to the domain of political action—have tended to argue for the permissibility of radical acts of disobedience without attending to the question of whether those who act permissibly in these ways are adequately differentiated from criminals.<sup>12</sup> To highlight the distinction, consider how a critic could say of such radical acts that even if they were permissible, those who are genuinely animated by conscientious convictions would not engage in them. This, as should be obvious, is a claim commonly made in public discourse. A further defense of how engaging in such acts *does not* impugn the conscientiousness of disobedients is thus also necessary.

#### 1. COMMUNICATIVE CONDITIONS

In this section, I briefly reconstruct the core components of Kimberley Brownlee's communicative principle of conscientiousness. This paves the way for the development of my account in the following two sections.

According to Brownlee's principle, genuine conviction has a communicative element. A disobedient who does not engage in such communication, and who remains silent, "necessarily casts doubt on the sincerity of [her] conviction."<sup>13</sup> A disobedient has reason to avoid inviting these doubts, for they may result in her being erroneously treated as a criminal, which often draws attention away from the issue against which she protests.<sup>14</sup> Brownlee's principle comprises four "communicative" conditions—consistency, universality, non-evasion, and dialogue. The communicative conditions are presented as individually necessary and jointly sufficient for someone to have conscientious conviction.<sup>15</sup> Taken together, the conditions specify and elaborate the general differentiation de-

12 For recent texts, see Brennan, *When All Else Fails*; Delmas, *A Duty to Resist*; Pasternak, "Political Rioting"; and Kapelner, "Revolution against Non-Violent Oppression." Elsewhere, I argue that activists have good reasons to engage in vandalism (Lim, "Vandalizing Tainted Commemorations").

13 Brownlee, *Conscience and Conviction*, 29.

14 See Terwindt, *When Protest Becomes Crime*, 234–36.

15 Brownlee has recharacterized these conditions in response to her critics. The non-evasion and dialogic conditions are now presented as corollaries of the consistency and universality conditions, respectively. This recharacterization does not affect my argument. See Brownlee, "Reply to Critics," 724.

mand.<sup>16</sup> When disobedients fail to satisfy the conditions, they accordingly fail to satisfy the differentiation demand.<sup>17</sup>

First, the consistency condition requires consistency among a disobedient's "judgements, motivations and conduct to the best extent that [she is] able."<sup>18</sup> Among other things, she avoids speech and conduct that contradicts her judgments or violates her commitments. Second, the universality condition requires disobedients to universalize their judgments. When they judge something to be *pro tanto* wrong, they must also judge it to be *pro tanto* wrong for others in similar circumstances—not just for themselves. For instance, a disobedient should not simply judge that it is wrong *for her* to participate in (an unjust) war but also that everyone who participates in (such a) war behaves wrongly.<sup>19</sup>

Third, the non-evasion condition requires disobedients to be willing to bear the risks of living in accordance with their convictions. They should not seek to evade the implications of their convictions, especially those arising from their disobedience. Here, Brownlee departs from John Rawls's famous specification of the condition—that individuals should willingly accept being arrested and facing legal punishment.<sup>20</sup> For Brownlee, all that is required is for disobedients to be willing to accept the *risk* of being arrested and punished.<sup>21</sup>

Finally, the dialogic condition requires disobedients to "be willing to communicate [their] conviction to others in an effort to engage them in reasoned deliberation about its merits."<sup>22</sup> When disobedients satisfy this condition, they treat others as reasoning agents rather than as those who may be (or are to be)

16 To preempt an exegetical worry: this characterization does not misunderstand Brownlee's argument. First, Brownlee explicitly presents the conditions as separating conscientious actors from "ordinary offenders [who] are not conscientiously motivated in any deep sense." Second, Brownlee is clear that conscientiousness is a *descriptive* property. Describing an individual as conscientious is not equivalent to making a moral evaluation of her person or actions. If so, the communicative conditions are not conditions for the *justification* or *permissibility* of actions. See Brownlee, *Conscience and Conviction*, 3–10, 17–18.

17 Here, I assume that Brownlee's principle is broadly plausible. Challenges to it are, of course, possible. For instance, we may think that conscientiousness requires critical and reflective endorsement such that genuine convictions are distinguished from those that are the result of ideological (or even propagandic) influences. Revisions to the principle do not affect my subsequent discussions.

18 Brownlee, *Conscience and Conviction*, 30.

19 Brownlee, *Conscience and Conviction*, 34–37. Satisfying the universality condition is, in principle, compatible with relying on the nonuniversalized claim to seek protections for one's conscientious refusal to participate in a given war.

20 Rawls, *A Theory of Justice*, 322.

21 Brownlee, *Conscience and Conviction*, 37–42.

22 Brownlee, *Conscience and Conviction*, 42.

coerced. Two caveats are important. First, the condition is not overly demanding—it does not require disobedients to *succeed* in communicating their conviction or persuading their audience; it requires only their willingness to do so. Second, a genuine dialogue is not mere assertion. Participants in a dialogue must be responsive to the possibility that they may be mistaken and ensure that their communication is likely to foster rather than detract from dialogue.<sup>23</sup>

The four conditions are broadly context sensitive. While it is important to satisfy them, doing so does not have conclusive weight. Depending on the context, other considerations—such as those to do with the “burdens of vulnerability, disadvantage, unpopularity, relative power, and the relative costs of communication”—may outweigh the requirement to behave in ways that satisfy the conditions.<sup>24</sup> In accommodating context sensitivity, Brownlee also accommodates the fact that disobedients may be committed to respecting and furthering other values.

According to Brownlee, three of the four conditions—consistency, non-evasion, and dialogue—have conative elements. They are connected to individuals’ actions. Because of these conative elements, the conditions are *practically testable*. Observers can look at the conduct of disobedients to check whether they satisfy the conditions. This is a more credible way of assessing whether disobedients have and act on the basis of conscientious convictions than simply taking their word for it.<sup>25</sup> A disobedient whose judgments, motivations, and actions are consistent shows that she genuinely has and acts on the basis of conscientious convictions. A disobedient who is non-evasive signals that her assertions about her convictions are not just talk. A disobedient who is willing to engage others in dialogue or to stand up for her convictions in a public way shows, again, the sincerity and seriousness of her convictions.<sup>26</sup> The universality condition does not have a conative element because it requires only universalized *pro tanto* judgments. Insofar as such judgments may be outweighed, the all-things-considered judgments on the basis of which individuals act may not reflect (and may indeed deviate from the demands of) their *pro tanto* judgments.

## 2. SHOWING CONSCIENTIOUSNESS

I begin the articulation of my account by highlighting a distinction, which is often missed in discussions of how disobedients are (or are to be) differentiated from criminals, between having conscientious convictions and showing that

23 Brownlee, *Conscience and Conviction*, 20, 42–44, 223.

24 Brownlee, *Conscience and Conviction*, 44.

25 Brownlee, *Conscience and Conviction*, 30.

26 Brownlee, *Conscience and Conviction*, 33–43.

one acts on the basis of such convictions. A brief discussion of the communicative conditions clarifies this distinction.

The communicative conditions may be characterized in two ways.<sup>27</sup> First, they may be understood as identifying and demarcating the category of those who *have* conscientious convictions. When disobedients fail to satisfy one or more of these conditions, they do not have conscientious convictions.<sup>28</sup> Call this *Possession*. Second, in satisfying the conditions, individuals *show* others that they act on the basis of conscientious convictions (though whether they have such convictions is, of course, a separate issue). Call this *Indication*. According to Indication, disobedients' satisfaction of the conditions assuages the doubts of observers about whether they act on the basis of conscientious convictions.<sup>29</sup> On both Possession and Indication, the conditions are necessary and sufficient.

Possession and Indication are intimately connected. Specifically, Indication succeeds because Possession sketches out a plausible view about what it is to have such convictions in the first place. The distinction and connection between Possession and Indication are crucial. However, Brownlee does not clearly distinguish them in articulating her account, nor does she offer an explicit discussion of their connection. They also appear to be missed by Brownlee's critics.<sup>30</sup>

Indeed, many critics of the communicative principle offer a structurally similar argument centering on the communicative conditions' susceptibility to a range of obvious counterexamples—comprising individuals who fail to satisfy one of the conditions yet who *actually have* conscientious convictions. For instance, Christopher Cowley offers the counterexample of a committed vegetarian who does not recognize the possibility of her being mistaken about whether eating meat is wrong. Cowley suggests that this case bears structural similarities to the activist Rosa Parks, who disobeyed laws mandating racial segregation. Parks is described as being similarly secure in her conviction that segregation is wrong. In both cases, the individuals are “not open to the possibility of error,”

27 The *characterization* of the conditions is distinct from their *specification*. The latter concerns what the conditions *pick out*. The former concerns how we understand the nature of the conditions themselves—including, among other things, their function, point, or more broadly, their relationship to the differentiation demand.

28 Brownlee, *Conscience and Conviction*, 33, 35, 38, 40, 43.

29 Brownlee, *Conscience and Conviction*, 31, 36–37, 38, 42.

30 Thomas E. Hill seems to be among the few critics who explicitly recognize that the communicative conditions can be characterized in different ways. He describes the conditions as those for an individual to have conscientious convictions (mapping on to Possession), and as identifying the convictions that the law should respect and protect. While I suspect that the conditions are not appropriately characterized in the latter way, I do not discuss this issue here. See Hill, “Conscientious Conviction and Conscience.”

yet are very plausibly described as having conscientious convictions.<sup>31</sup> This is a challenge to Brownlee's claim that such recognition is among the necessary conditions for having conscientious convictions.<sup>32</sup> Similarly, Candice Delmas provides several counterexamples of individuals who run afoul of the "narrow conception" of conscientiousness that Brownlee sketches yet who actually have conscientious convictions. According to Delmas, a Catholic who engages, before marriage, in various sorts of intimate physical contact short of coitus—thus failing the consistency condition—may nonetheless be conscientiously devout. Someone who fails to satisfy the universality condition in her refusal to pass judgment on women who underwent or plan to undergo abortions may nonetheless have conscientious convictions against abortion.<sup>33</sup> Cowley and Delmas are not alone; similar arguments are made elsewhere.<sup>34</sup>

These counterexamples are intended to support the position that the communicative conditions are specified too stringently or narrowly and thus deliver extensionally inaccurate verdicts about who has conscientious convictions. If successful—if, that is, the agents to which the counterexamples refer can plausibly be said to have conscientious convictions—they pose a serious challenge to the characterization of the conditions as necessary and sufficient for individuals to have conscientious convictions.

31 Cowley, "Conscientious Objection and the Limits of Dialogue," 1009.

32 Brownlee, *Conscience and Conviction*, 2019. In response to critics, Brownlee appears to weaken the specification of, and the requirement imposed by, the dialogic condition. Someone like Rosa Parks would count as conscientious if she "would have tried to understand" her opponents' motivations and commitments and if she "would also have sought to see things" from their perspectives. This revision is, however, also susceptible to the kinds of counterexamples raised by her critics; I set it aside (Brownlee, "Reply to Critics," 728).

33 Delmas also lists, among her examples, the case of an individual raised in a very conservative environment who "might be evasive and non-dialogic as she comes to shed her parents' and peers' views and develops liberal conscientious convictions." She credits Alon Harel as the inspiration for this example. However, this example does not challenge the specification of the communicative conditions. Since the individual in concern has not fully developed her views, she may plausibly (especially at the start) be described as not (yet) having conscientious convictions. Of course, this does not mean that these beliefs are insignificant—they are preconditions of or precursors to conscientious convictions. We may even decide that these beliefs that fall short of the standard of conscientiousness are significant enough to be included within the protections or exceptions accorded to conscientious convictions. This, however, is not the same as saying that the standard of conscientiousness *itself*—which determines what counts as *having conscientious convictions*—is overly narrow and should be relaxed. See Delmas, "False Convictions and True Conscience," 409–10; and Harel, review of *Conscience and Conviction*.

34 Among others, see Smith, "The Burdens of Conviction," 694–97; and Coady, review of *Conscience and Conviction*, 502–3.

While the counterexamples are intuitively plausible, the critics are mistaken in specifying the challenge they pose. Consider their claim that even though the agents in the counterexamples fail to satisfy one of the communicative conditions, they nonetheless *actually have* conscientious convictions. Whether these agents actually have conscientious convictions, however, is beside the point. Recall that our task concerns differentiating disobedients while lacking unmediated access to their internal states. The differentiation is to be secured from the perspective of their audience. From this perspective, an individual fails the differentiation demand when she fails to behave in ways that show others that she acts on the basis of conscientious convictions—even though she may actually have such convictions.

Given this, we should understand the counterexamples differently. They show that disobedients may succeed in *showing* others that they act on the basis of conscientious convictions despite their failure to satisfy one of the communicative conditions. This is because their actions satisfy *enough* of the (other) conditions such that observers may securely and confidently judge them as acting on the basis of conscientious convictions. The fact that we indeed make such judgments—and often confidently so—is, I take it, delivered by the critics' counterexamples. The counterexamples, then, are rightly a challenge to Indication, but not to Possession.

We may respond to this challenge to Indication in two ways. First, we may re-specify the conditions so that they accurately pick out those who show that they act on the basis of conscientious convictions. That is, we may try to ensure that they are indeed the right necessary and sufficient conditions. While this may be a plausible option, it is not one I take here. I suggest that we choose, instead, to recharacterize the conditions. Following from earlier discussions, the most natural recharacterization of the conditions is that they delineate what is *paradigmatically* or typically true of those who successfully show others that they act on the basis of conscientious conviction—rather than what is *necessary and sufficient* for doing so. That is, we should understand Indication as outlining the paradigmatic conditions for showing others that one acts on the basis of conscientious convictions. Call this *Indication\**. This is the first core aspect of my account. It responds to the challenge posed by the counterexamples not by weakening the specification of the conditions but by weakening how we characterize them.

There are at least two reasons to endorse *Indication\**. First, it allows us to sidestep the trade of counterexamples that is invited and facilitated by characterizing the conditions as necessary and sufficient. According to *Indication\**, the failure of an individual to satisfy one of the communicative conditions (or perhaps to meet all of them fully) does not automatically mean that she fails to show others that she acts on the basis of conscientious convictions. To reach that verdict, we must pay attention to the specifics. Among other things, we would have

to examine whether she has satisfied the other conditions, and by how much; we would also have to engage in deliberations about whether her satisfaction of those other conditions allows us to judge her as having shown that she acts on the basis of conscientious convictions. This characterization, then, facilitates more nuanced discussions—an especially important payoff when we encounter borderline or vague cases. Second, Indication\* is accommodating of imperfect beings like ourselves. Acting in ways that fail to satisfy just one of the conditions does not automatically mean that we fail to show others that we act on the basis of conscientious convictions. And insofar as being regarded as such is important for the protections or exceptions that are typically granted to conscientious disobedients, Indication\* is more humane than Indication.<sup>35</sup>

Adopting Indication\*, however, results in some vagueness in determining whether disobedients succeed in showing others that they act on the basis of conscientious convictions. In some borderline circumstances, Indication\* may not even deliver any determinate answers. Here, my response is resolute—we should accommodate rather than eliminate this vagueness. We should not try to make our determination of who succeeds in showing others that they act on the basis of conscientious convictions seem clearer than it actually is. Doing so is concordant with the initial problem—that we do not have direct and unmediated access to the internal states of the actors whose actions we are tasked with evaluating. Moreover, given that we can confidently judge whether a disobedient acts on the basis of conscientious convictions despite her failure to satisfy every single one of the conditions, there appears to be little practical payoff in construing the conditions as necessary and sufficient.

### 3. HOLISTIC ASSESSMENT

According to Indication\*, a disobedient who fails to satisfy one of the communicative conditions can nonetheless succeed in showing others that she acts on the basis of conscientious convictions. This leaves open the possibility that when such a disobedient engages in an act that fails to satisfy *most* or *all* of the conditions, she fails to show that she acts on the basis of those convictions—thus failing to satisfy the differentiation demand. In this section, I argue that this possibility should not bother us too much—in many cases, disobedients can

35 We might think that the reasons for endorsing Indication\* also support a similar recharacterization of Possession—such that the conditions are paradigmatically satisfied by those who have conscientious convictions, rather than necessary and sufficient conditions for having such convictions. I do not take a stance on this issue here. I leave open the possibility that the conditions are indeed necessary and sufficient for having conscientious convictions.

satisfy the differentiation demand even though their singular or occasional acts of disobedience may fail to satisfy most or even all of the conditions. This is the second core aspect of my account.

Consider a committed environmentalist, Aly, who dedicates a significant portion of her life to campaigning and activism and who generally behaves in ways that satisfy the communicative conditions in most areas of her life and political activity. On one occasion, she anonymously dumps pollutants into the waters at a beach that is much loved and frequented by the locals, intending to draw attention to and protest pollution and environmental degradation. Aly's action appears to violate all the communicative conditions. Her polluting act appears to be inconsistent with her stance against pollution. It seems to violate the universality condition insofar as she does something she thinks is wrong for others to do. The anonymity of her action clearly violates the non-evasion condition and plausibly also the dialogic condition. It appears that, in this case, Aly fails to show that she acts on the basis of conscientious convictions—and thus fails to differentiate herself from criminals.

Two related—and increasingly resolute—responses to this are available for the case of Aly (and those similar to it). First, we can challenge the claim that Aly's act violates all the communicative conditions. If this succeeds, we mitigate or even eliminate the doubts about whether she acts on the basis of conscientious convictions. This first response turns on a finer-grained description of the act. Aly may be concerned not about pollution and environmental degradation *simpliciter* but with the unequal distribution of burdens imposed by pollution and environmental degradation. She may decry the fact that the burdens of pollution and degradation are disproportionately borne by the most disadvantaged individuals in society (or in the world). This, as it turns out, is one of the most common claims made by environmental activists and disobedients. Aly's convictions may be more accurately and plausibly presented when they are understood as “fine grained” or specific, rather than “coarse grained” or general. If so, her act of polluting the beach—assuming it does not contribute to the pollution and degradation affecting the most disadvantaged—is concordant with what she stands for.<sup>36</sup> She does not necessarily violate the consistency or universality conditions. Even if she remains anonymous to the public, the activist communi-

36 There are complications arising from indirect contribution—especially where and how to establish the threshold beyond which indirect contributions are to be considered as part of an act and thus feature in our descriptions and evaluations of the latter. These complications are not unique to my account; they are faced by accounts of disobedience more generally. They concern how we should set a threshold between two implausible extremes—one where *all* indirect contributions count and the other where *no* such contributions count. I do not address these complications here.



ty of which she is part may take responsibility for her actions. If so, the act may not fail the non-evasion condition.

This response—which centers on how we describe disobedients' convictions—is often generalizable.<sup>37</sup> One frequent criticism of disobedient acts—especially those that impose obvious and significant burdens on others—is that they represent everything the disobedients purport to stand against. For instance, disobedients whose actions result in property damage or harm to others are often described as behaving in ways that violate their convictions about protecting the interests of, or avoiding harm to, others—failing the consistency condition. They are often also described as granting themselves the license to engage in actions of the kind they protest—failing to satisfy the universality condition. In these cases, disobedients are often denounced as criminals on the basis of their engagement in these actions, based on the assumption that those whose disobedient acts are undergirded by conscientious convictions would not engage in them. Now, we see that the success of these criticisms in threatening the differentiation of disobedients from criminals actually turns on the unstated assumption that the commitments of the disobedients are most plausibly understood in general rather than specific terms. In many cases, this assumption is unwarranted. Among other things, even disobedients who participate in arson, rioting, vandalism, or vigilantism—actions that are most frequently regarded as mere criminality—can be understood as behaving in ways that are concordant with their concerns and commitments. Their destructive actions are typically neither random nor wanton, but directed at those who are complicit in bringing about the injustice against which they protest.<sup>38</sup> Similar arguments have also been marshaled in defense of certain forms of vandalism and vigilantism.<sup>39</sup>

In sum, Aly's act (and those similar to it) may not actually violate most or all of the communicative conditions. If so—and drawing from our earlier discussions of Indication\*—she may still succeed in showing that she acts on the basis of conscientious convictions and in differentiating herself from criminals. The first response reiterates the earlier caution: in describing and evaluating disobedients and their actions, we must pay attention to the specifics. However, this response leaves open the possibility of cases where the disobedient act in concern

37 For a recent and extensive study of how the interests of various groups and actors—including public officials, businesses, prosecutors, and other citizens—shape whether and how acts of dissent (including disobedience) are *described* as criminal activities, see Terwindt, *When Protest Becomes Crime*.

38 For representative texts, see Fogelson, "Violence and Grievances"; Waddington, "The Madness of the Mob?"; and Moran and Waddington, *Riots*.

39 Brennan, *When All Else Fails*; Delmas, *A Duty to Resist*.

*actually* violates most or all of the communicative conditions. It appears that in committing such acts, the disobedient is undifferentiated from criminals. A further response is needed.

It may seem that a simple appeal to the broadly context-sensitive nature of the communicative conditions can help us to defend Aly's differentiation from criminals, even if she behaves in a way that violates most or all of the conditions. Recall that the satisfaction of the conditions does not have conclusive weight, especially relative to other considerations that may apply to any given disobedient. If so, the accommodation of context sensitivity means that the failure of a disobedient to meet most or even all of the communicative conditions *does not necessarily* indicate her *lack* of conscientious convictions. The appeal to context sensitivity is a resolute response insofar as it denies that we can make obvious or easy inferences that Aly lacks conscientious convictions on the basis of her actions that fail to satisfy most or all of the communicative conditions.

However, this response falls short of defending her. As we have seen earlier, whether a disobedient *actually has* conscientious convictions is beside the point. The demand facing disobedients is to differentiate themselves from criminals, given that observers do not have unmediated access to their mental states and thus need more than their mere assertions as reassurance that their disobedience is undergirded by their conscientious convictions. The reliance on their behavior is, in a sense, all that we have. The appeal to context sensitivity fails precisely where it is needed—it does not secure the claim that in violating most or all of the communicative conditions, Aly successfully *shows* others that she acts on the basis of conscientious convictions. Moreover, in accommodating the possibility that disobedients may not (and need not) behave in ways that satisfy the communicative conditions—which shows others that they act on the basis of conscientious convictions—we seem to return to the initial problem of having to rely on their assertions. In the context of Indication and Indication\*, the specter is raised that the incorporation of context sensitivity “defeats the point of the communicative principle of conscientious conviction, which is to guarantee that the sincerity of our commitments be visible to all, and that no doubt be cast on it.”<sup>40</sup> A simple appeal to context sensitivity here may render the practical tests associated with satisfying the conditions pointless.<sup>41</sup> I set it aside.

40 Delmas, “False Convictions and True Conscience,” 411. Delmas does not appear to distinguish between Possession and Indication. My claim here is that the worry is apt for the latter and not the former.

41 This means that there is an unresolved internal inconsistency in Brownlee's account of the communicative principle of conscientiousness concerning Indication (but not Possession). Since my concerns are not exegetical, I set this issue aside.

My second resolute response to the challenge posed by Aly's case begins from the recognition that the communicative conditions are a specification and elaboration of the more general differentiation demand. We are concerned with whether disobedients satisfy the conditions only because that is a way for us to determine if disobedients satisfy the differentiation demand. In recognizing this, however, we glimpse the possibility that a disobedient's satisfaction of the communicative conditions may come apart from her satisfaction of the differentiation demand. That is, she may fail at the former yet succeed in the latter. This, I suggest, is precisely so in the case of Aly.

To secure this claim, we need to adopt a holistic assessment of Aly and her act that violates most or all of the communicative conditions. Call this *Holism*. This is the second core aspect of my account. Holism covers two domains—the assessment of individuals and their actions. I discuss these in turn.

First, in our assessments of disobedients, we should look at their past and ongoing behavior in most or all areas of their lives, rather than simply looking at singular acts.<sup>42</sup> In the case of Aly, we should not narrowly focus on the act that violates most or all of the communicative conditions. Instead, we should attempt to get a fuller view of her actions prior to and concurrent with that act. Here, we see that she has generally behaved in ways that satisfy the conditions in most other areas of her life and political activity. She has generally (and perhaps even plentifully) shown herself to act on the basis of conscientious convictions. This licenses our inference that even though Aly behaves in a way that fails to satisfy most or all of the communicative conditions *in this case*, she nonetheless acts on the basis of conscientious convictions. If this inference succeeds, her failure in this case does not threaten or eliminate her differentiation from criminals. More generally, disobedients can satisfy the differentiation demand even though they do not *always* behave in ways that satisfy most or all of the communicative conditions.

The holistic assessment of disobedients is intuitively plausible. We are typically reluctant to judge activists who have spent large parts of their lives acting on the basis of conscientious convictions as undifferentiated from criminals—those who do not act on the basis of such convictions—simply on the basis of their engaging in a single act that violates most or all of the communicative conditions. In practical terms, we typically can securely differentiate even a disobedient who engages in radical actions (such as arson, rioting, vandalism, or vigilantism) from criminals who do not act on the basis of conscientious convictions. The intuitive plausibility of holistic assessment extends to cases that are the reverse of Aly's.

42 Holism may even accommodate the *retrospective* differentiation of disobedients' past actions, on the basis of their subsequent actions or revelations. This possibility gives rise to complications that I do not address here.

Consider a lawbreaker who has *not* shown that she acts on the basis of conscientious convictions in most or all areas of her life, but who, on one occasion, breaks the law in a way that satisfies all of the communicative conditions. She purportedly shows that in this case she acts on the basis of conscientious convictions. Despite this, we are likely to have serious suspicions about whether she *truly* acts on the basis of conscientious convictions. These suspicions, however, are unwarranted (and perhaps even unintelligible) if we focus only on the act itself. They are warranted only if we assess the lawbreaker holistically.

On a holistic assessment, disobedients' differentiation from criminals is threatened or even eliminated only when they have *rarely or never* behaved in ways that satisfy the communicative conditions, and thus rarely or never given any indications of their conscientious convictions prior to the act in concern (which fails to satisfy most or all of the communicative conditions). In such cases, however, it may not be problematic for us to adhere, *even dogmatically*, to the verdict delivered—that the disobedients in concern are not adequately differentiated (if at all) from criminals. Of course, this verdict may be mistaken—the disobedients may *actually* be acting on the basis of their conscientious convictions. But given that they have rarely or never behaved in ways that indicate their conscientious convictions—along with our having no unmediated access to their mental states and with our interests in not being strung along—even our dogmatism may be appropriate. Beyond these extreme and clear-cut cases, we would have to engage in nuanced discussions—concerning whether (and the extent to which) disobedients satisfy the other communicative conditions in other areas of their lives. Holism facilitates more nuanced discussions than an account that judges disobedients to be undifferentiated from criminals on the basis of singular acts that fail to satisfy most or all of the conditions. While this point may seem obvious, it is not often heeded in public and philosophical discourse.<sup>43</sup> Commentators are often quick to denounce disobedients as criminals based on one or a few of the actions that they engage in—disregarding the relevance and significance of their previous behavior.

The main upshot of a holistic assessment of individuals is that it reveals that our concern with acts of disobedience that violate most or all of the communicative conditions really matters only in borderline cases, where the violations threaten or eliminate the differentiation between a disobedient and a criminal. In many cases, whether any given act of disobedience violates most or all of the conditions is likely to be immaterial for the purposes of differentiation. We

43 This raises questions about the importance of the character or settled dispositions of disobedients for determining the seriousness or authenticity of their actions. I do not address them here.

should not exaggerate the importance of any given act of disobedience for the differentiation of disobedients from criminals.

The second domain that Holism covers concerns acts of disobedience. Without this, Holism would be implausibly permissive. This is because the holistic assessment of disobedients does not appear to set restrictions on what exactly disobedients can do. Here, the worry is that some acts that fail to satisfy most or all of the communicative conditions are worse than others. An act of pollution intuitively appears to be significantly different from one of terror bombing in terms of how they affect the differentiation of their perpetrators from criminals—even if they both similarly violate most or all of the conditions.

Here, we should look at whether a given act of disobedience can be understood as part of a disobedient's broader project to live in accordance with her convictions, or to bring about conditions in which she can do so. It is only when the act cannot be understood in this way—when it has no discernible connections to any broader projects—that the worry about the disobedient's differentiation has bite. Otherwise, the radical nature of the act does not threaten or eliminate the disobedient's differentiation from criminals.

Holism helps us to accommodate acts of disobedience that are radical *because* of the contexts in which disobedients find themselves.<sup>44</sup> Disobedients' deliberations about their political action (including, but not limited to, disobedience) are complicated. They have to weigh up a range of complex and interconnected considerations salient in the contexts in which they operate. Among other things, these may pertain to the relationship between the disobedients and their audience, the expected reactions of other citizens and public officials, the expected responses from and implications for other activist groups, the organization of the media, the possibility of publicizing their act without distortion, and so on. Disobedients may find themselves in a society that has repeatedly ignored or dismissed their previous acts of legal protests or their constrained disobedience. Or they may find themselves in a society plagued by injustices so severe and urgent that legal protests or constrained disobedience would be pointless. In these contexts, disobedients may have no other choice than to engage in what are regarded as radical acts if they wish to make any advances in their broader projects. According to Holism, as long as these acts can be seen as part of disobedients' broader projects, the fact that they are radical neither threatens nor eliminates the disobedients' differentiation from criminals.

A simplified example (which centers on the relationship between an actor

44 For a discussion of how some prosecutors and legal systems describe and evaluate disobedient actions in a way that ignores the broader context in which they are carried out, and the problems accompanying that approach, see Terwindt, *When Protest Becomes Crime*, 56–65.

and her audience) illustrates the point. Suppose that Betty has just discovered that her housemate, Charlie, has stolen money from her to pay for recreational drugs and that this is not the first time that Charlie has done this. Betty has tried, by having conversations with him on several occasions, to gently persuade Charlie to stop. Suppose that Betty now breaks into Charlie's room and steals his belongings—to teach him a lesson about the impact and invasiveness of theft. In her anger, she also damages some of Charlie's belongings. We can plausibly suppose that Betty's actions, while understandable, are drastic and radical. In this case, would we say of Betty that she is therefore undifferentiated from a criminal who behaves similarly? I do not believe so. Betty stands in a relationship to Charlie and has tried to engage Charlie in milder ways. Her current actions, while drastic, make sense as a constituent of a bigger project—of getting Charlie to change his behavior. The same cannot be said for the criminal who breaks into Charlie's room and does exactly the same things. This point is extendable to the case of disobedients who engage in radical acts of disobedience.

Two clarifications are in order. First, my claim is that disobedients who engage in radical acts are not, simply on the basis that those acts are radical, therefore *undifferentiated* from criminals. I do not claim here that such radical acts are *justified* or *permissible*. We may have several serious reservations about Betty's actions and, by extension, those of disobedients who engage in radical acts. These may have to do with the fact that they may be unnecessary, disproportionate, or aimed at the wrong targets, among others. Indeed, our recognition that disobedients are securely differentiated from criminals is *compatible* with our directing harsh criticisms at, or flatly rejecting, their actions. What we cannot do, however, is to criticize and reject those actions *on the basis that* their radical characters render the disobedients no different from criminals.

Second, we must not adopt an overly restrictive view of whether and how an act can be seen as part of a disobedient's broader project. Radical acts are often viewed negatively. Among other things, they are criticized for being ineffective for achieving the disobedient's goals or counterproductive to the achievement of those goals. It may be thought that ineffective and counterproductive actions cannot be understood as part of a disobedient's broader project. This, however, is a mistake. Acts of disobedience that are often touted as ineffective or counterproductive may actually serve goals that critics have missed.<sup>45</sup> Among other things, disobedients are also interested in generating publicity, starting discussions, sparking and sustaining interests, building solidarity with other activists,

45 It has been argued that radical acts are not necessarily ineffective or counterproductive. For further discussion (though under the classification of "uncivil" acts), see Delmas, *A Duty to Resist*, 58–67.

or simply challenging everyday practices.<sup>46</sup> An action that is ineffective or counterproductive with regard to one goal may instead serve other goals. We must be cautious on two issues. We must not assume that one goal (or a subset of them) is the most fundamental, and the only goal against which the assessments of the effectiveness or counterproductivity of acts are to be made. Moreover, even if there were indeed one fundamental goal—and one that is endorsed by the disobedients themselves—we must not assume that disobedients must always act in ways that directly and immediately serve that goal.

The discussions in this section pose a serious challenge to the views—common in much philosophical and public discourse—that the radical quality of certain actions *on its own* throws doubt on whether the disobedients in concern act on the basis of conscientious convictions, and that radical acts automatically render their perpetrators equivalent to criminals. Actions that are typically regarded as radical in these ways include arson, rioting, vandalism, or vigilantism, among others. On my account, we see that these views are critically incomplete. We must not fixate on the radical quality of the actions. We must also look at the actors and the context in which the actions are situated. Holism provides a conditional defense of radical action. Insofar as the disobedients who engage in radical actions have generally shown that they act on the basis of their conscientious convictions, and insofar as their act can be seen as part of a broader project to which they are committed, the radical character of their act does not threaten or eliminate their differentiation from criminals. It is only when they fail either of the conditions that their differentiation is threatened.<sup>47</sup> In practical terms, adopting Holism means that disobedients are securely differentiated in a broad range of cases—even for radical acts that are typically equated with criminality.

#### 4. CONSTRAINING DISOBEDIENTS

In the preceding sections, I have argued that we should not narrowly focus on whether acts of disobedience satisfy the communicative conditions. Instead, we should also examine whether those acts can be understood as part of disobedients' broader projects and whether the disobedients have provided adequate indication that they act on the basis of conscientious convictions. In this section, I consider two related objections, both of which center on the worry that my account imposes overly stringent constraints on disobedients.

My defense of acts of disobedience that fail to satisfy most or all of the

<sup>46</sup> Walzer, *Political Action*.

<sup>47</sup> There is a further question about the relationship between the two conditions—namely, which is more fundamental—that I set aside for future work.

communicative conditions rests on the claim that insofar as disobedients have provided ample indication that they act (or have acted) on the basis of conscientious convictions in other areas of their lives, their failure to satisfy the conditions in any given case does not threaten or eliminate their differentiation from criminals. This may be thought to entail a severe constraint on disobedients. They have to ensure that their past and ongoing behavior in other areas of their lives—leading up to their acts of disobedience—provides adequate indication that they have and act on the basis of conscientious convictions. Not just anyone can engage in lawbreaking political action while simultaneously satisfying the differentiation demand. In many cases, we may find that disobedients have failed to meet this severe constraint, and thus that my account would unfairly treat them as undifferentiated from criminals. For example, individuals who are not ordinarily social justice activists appear to be undifferentiated from criminals if they participate in large-scale disobedience protesting injustice.

Several considerations mitigate the apparent severity of this constraint. First, the differentiation demand is not the most important demand on political disobedience or action. In some circumstances, satisfying the differentiation demand may take a back seat to the need to mitigate or eliminate serious injustices. In these cases, there may be good reason for individuals to behave disobediently, even in the knowledge that they may not be differentiated from criminals.

Second, what we are asking of disobedients is simply that they show others, in other areas of their lives beyond their disobedience, that they indeed have and act on the basis of their own conscientious convictions. These are convictions that—if disobedients indeed possess them—are among the most fundamental and central in their importance to the disobedients in concern.<sup>48</sup> Individuals with such convictions aim, typically, to behave in ways that preserve their integrity.<sup>49</sup> Given this, even the requirement that they live in accordance with their deepest convictions, as far and as often as is possible, does not strike me as implausible. Of course, we must make room for context sensitivity in our assessments of whether any given disobedient has behaved in this way in other aspects of her life. We should take seriously and try to accommodate the possibilities, discussed earlier, that in some circumstances the burdens associated with living in accordance with one's conscientious convictions may be too onerous.

That being said, and third, we should take an ecumenical view of what counts as doing enough to indicate one's conscientious convictions. Insofar as conscientious convictions are deeply important to individuals—including, among other things, structuring how those individuals see and interact with the world—we

48 Brownlee, *Conscience and Conviction*, 7.

49 McCloskey, "Conscientious Disobedience of the Law," 537.



should and can expect that they are shown even through these individuals' mundane and ordinary actions. More concretely, we can say that activities such as the following should be regarded as providing indication that an individual acts on the basis of her conscientious convictions: participating in or contributing to activist organizations or events, attending teach-ins or other lectures, attempting to engage in conversations with or convert others to one's cause (which can include, at the seemingly most trivial, conversations on social media), and signing and circulating petitions, among others. We should neither think nor require that disobedients only indicate their conscientious convictions when they engage in high-stakes activities.<sup>50</sup>

While ecumenical, this constraint is not toothless. Among other things, individuals who participate in lawbreaking activity on impulse without having provided adequate indication of their conscientiousness in other aspects of their lives would be judged as insecurely differentiated (if at all) from criminals. In these cases, we want them to do more than they have done to show (or show more clearly) that they indeed have and act on the basis of conscientious convictions.

The second objection is that my account delivers the wrong result in the cases of disobedients who have recently changed their conscientious convictions—"converts"—and who act on the basis of their newfound convictions. In such cases, my account appears to incorrectly judge—especially on the basis of their *previous* behavior—that converts are not conscientious. Consider, for instance, an individual who has not previously behaved in any ways that indicate her conscientious convictions. Suppose that she now possesses strong conscientious convictions about police brutality and violence and proceeds to engage in a radical act of disobedience (perhaps anonymously setting fire to a police station). Here, it seems that on the basis of Holism she faces insurmountable difficulties in showing others that she does act on the basis of her conscientious convictions. She is wrongly picked out as undifferentiated from a criminal, when in fact she has and acts on the basis of conscientious convictions.

50 Iris Murdoch puts the point beautifully:

When we apprehend and assess other people we do not consider only their solutions to specifiable practical problems, we consider something more elusive which may be called their total vision of life, as shown in their mode of speech or silence, their choice of words, their assessments of others, their conception of their own lives, what they think attractive or praise-worthy, what they think funny: in short, the configurations of their thought which shows continually in their reactions and conversation. These things, which may be overtly and comprehensibly displayed or inwardly elaborated and guessed at, constitute what . . . one may call the texture of a man's being or the nature of his personal vision. (Hepburn and Murdoch, "Symposium," 39)

While this is a serious problem, two considerations mitigate its sting. First, changes in deep convictions do not usually happen overnight. They typically take time to develop. Their development is also typically facilitated by interactions and conversations with others. Conscientious convictions do not, as it were, “leap as Athena did full-grown and armed from Zeus’s head.”<sup>51</sup> If so, there may well be plenty of opportunities for the convert to behave in ways that indicate her changed (or changing) convictions. Taken together with the ecumenical view of what counts as doing enough to indicate one’s conscientious convictions, this allows my account to pick out even fairly recent converts.<sup>52</sup> Second, if such opportunities are rare, we may reasonably be resolute in our skepticism that the convert has and acts on the basis of her conscientious convictions and in our belief that she may not be adequately differentiated (if at all) from criminals. This is connected to the very problem we began with that motivates the differentiation demand—that we do not have unmediated access to others’ internal states and that we have interests in avoiding being strung along. While this resoluteness is unsatisfactory—it does not *guarantee* extensional accuracy—it appears to be a more general problem with attempting to secure differentiation from the perspective of one’s audience. The hope, then, is that in light of the first consideration, the circumstances in which we have to stand resolute are few and far between.

## 5. CONCLUSION

Conscientious disobedients face the demand to differentiate themselves from criminals whose actions also violate laws but are not undergirded by conscientious convictions. In public and philosophical discourse—though with different levels of sophistication—conscientious disobedients are often criticized on the basis that their actions render them no different from criminals. In this essay, I have argued otherwise, by articulating and defending the claim that disobedients who engage in radical acts of disobedience may still be securely differentiated from criminals. The category of conscientious disobedients is potentially more inclusive than has been commonly assumed within public and philosophical discourse. We have reason to revise our judgments and treatments of such disobedients.

51 Harel, review of *Conscience and Conviction*.

52 This might not be enough for us to accurately pick out those individuals who in fact have sudden and radical transformations in their conscientious convictions—perhaps and especially of a religious sort. Here, and in light of how the differentiation demand has been specified, I am inclined to take the resolute response I discuss presently. Alternative ways of specifying the differentiation—such as that given by Arendt in *Crises of the Republic*—may well avoid this problem, albeit incurring other costs.

The main upshot of my account is that it significantly reframes what the salient questions are when we are confronted with disobedience. Rather than narrowly focus on acts of disobedience that are committed, we must also direct our attention to the disobedients themselves and the contexts in which the acts are committed. On my account, the differentiation demand can be satisfied even by disobedients who engage in what are typically regarded as radical actions. In practical terms, this means that even disobedients who engage in actions such as arson, rioting, vandalism, or vigilantism can also successfully differentiate themselves from criminals. We should be much less eager to denounce conscientious disobedients on the basis of their engagement in this or that action, and we must pay more attention to other aspects of their lives and actions. I believe that this account best vindicates the claims—made by many disobedients—that even their radical actions are undergirded by conscientious convictions and that they are not criminals.<sup>53</sup>

*Stanford University and Nanyang Technological University*  
cm.lim@ntu.edu.sg

#### REFERENCES

- Arendt, Hannah. *Crises of the Republic*. New York: Harcourt Brace, 1969.
- Beinart, Peter. “Left Wing Protests Are Crossing the Line.” *The Atlantic*, November 16, 2018. <https://www.theatlantic.com/ideas/archive/2018/11/protests-tucker-carlsons-home-crossed-line/576001>.
- Bejan, Teresa M. *Mere Civility*. Cambridge, MA: Harvard University Press, 2017.
- Brennan, Jason. *When All Else Fails: The Ethics of Resistance to State Injustice*. Princeton: Princeton University Press, 2018.
- Brownlee, Kimberley. *Conscience and Conviction: The Case for Civil Disobedience*. Oxford: Oxford University Press, 2012.
- . “Reply to Critics.” *Criminal Law and Philosophy* 10, no. 4 (December 2016): 721–39.
- Celikates, Robin. “Rethinking Civil Disobedience as a Practice of Contesta-

53 I am grateful for comments from and discussions with Yuna Blajer de la Garza, Kimberley Brownlee, Daniel Butt, Kathleen Creel, Cécile Fabre, Kailing Fu, Abby Jaques, Dongxian Jiang, Todd Karhu, Geoff Keeling, Nannerl Keohane, Desiree Lim, Anne Newman, Avia Pasternak, Joseph Raz, Rob Reich, Thomas Simpson, Alicia Steinmetz, Zofia Stemplowska, Lee Wilson, Jonathan Wolff, the audiences at Ockham Society (Oxford, 2018), and two anonymous reviewers for this journal.

- tion—Beyond the Liberal Paradigm.” *Constellations* 23, no. 1 (March 2016): 37–45.
- Coady, C. A. J. Review of *Conscience and Conviction: The Case for Civil Disobedience*, by Kimberley Brownlee. *Journal of Value Inquiry* 50, no. 2 (June 2016): 501–6.
- Cochrane, Emily. “Sarah Huckabee Sanders Was Asked to Leave Restaurant over White House Work.” *New York Times*, June 23, 2018. <https://www.nytimes.com/2018/06/23/us/politics/sarah-huckabee-sanders-restaurant.html>.
- Cowley, Christopher. “Conscientious Objection and the Limits of Dialogue.” *Philosophy and Social Criticism* 42, no. 10 (December 2016): 1004–14.
- Davis, Angela. “Race and Criminalization: Black Americans and the Punishment Industry.” In *The House That Race Built: Original Essays by Toni Morrison, Angela Y. Davis, Cornel West, and Others on Black Americans and Politics in America Today*, edited by Wahneema Lubiano, 264–79. New York: Random House, 1997.
- Delmas, Candice. *A Duty to Resist: When Disobedience Should Be Uncivil*. Oxford: Oxford University Press, 2018.
- . “False Convictions and True Conscience.” *Oxford Journal of Legal Studies* 35, no. 2 (Summer 2015): 403–25.
- Estlund, David M. “Deliberation Down and Dirty: Must Political Expression Be Civil?” In *The Boundaries of Freedom of Expression and Order in American Democracy*, edited by Thomas R. Hensley, 49–67. Kent, OH: Kent State University Press, 2001.
- Fogelson, Robert M. “Violence and Grievances: Reflections on the 1960s Riots.” *Journal of Social Issues* 26, no. 1 (Winter 1970): 141–63.
- Harcourt, Bernard. “The Politics of Incivility.” *Arizona Law Review* 54 (January 2012): 345–73.
- Harel, Alon. Review of *Conscience and Conviction: The Case for Civil Disobedience*, by Kimberley Brownlee. *Notre Dame Philosophical Reviews*. February 29, 2013. <https://ndpr.nd.edu/reviews/conscience-and-conviction-the-case-for-civil-disobedience/>.
- Hepburn, R. W., and Iris Murdoch. “Symposium: Vision and Choice in Morality.” *Aristotelian Society Supplementary Volume* 30 (1956): 14–58.
- Hill, Thomas E. “Conscientious Conviction and Conscience.” *Criminal Law and Philosophy* 10, no. 4 (December 2016): 677–92.
- Kapelner, Zsolt. “Revolution against Non-Violent Oppression.” *Res Publica* 25, no. 4 (November 2019): 445–61.
- Lim, Chong-Ming. “Vandalizing Tainted Commemorations.” *Philosophy and Public Affairs* 48, no. 2 (Spring 2020): 185–216.

- Lovell, Jarret S. *Crimes of Dissent: Civil Disobedience, Criminal Justice, and the Politics of Conscience*. New York: New York University Press, 2009.
- McCloskey, H. J. "Conscientious Disobedience of the Law: Its Necessity, Justification, and Problems to Which It Gives Rise." *Philosophy and Phenomenological Research* 40, no. 4 (June 1980): 536–57.
- Moran, Matthew, and David P. Waddington. *Riots: An International Comparison*. London: Palgrave Macmillan, 2016.
- Osterweil, Vicky. *In Defense of Looting: A Riotous History of Uncivil Action*. New York: Bold Type Books, 2019.
- Pasternak, Avia. "Political Rioting: A Moral Assessment." *Philosophy and Public Affairs* 46, no. 4 (Fall 2018): 384–418.
- Rawls, John. *A Theory of Justice*. Rev. ed. Cambridge: Belknap Press, 1999.
- Sarat, Austin, ed. *Civility, Legality, and Justice in America*. Cambridge: Cambridge University Press, 2014.
- Scott, James C. *Weapons of the Weak: Everyday Forms of Peasant Resistance*. New Haven: Yale University Press, 1985.
- Smith, William. "The Burdens of Conviction: Brownlee on Civil Disobedience." *Criminal Law and Philosophy* 10, no. 4 (December 2016): 693–706.
- Swaine, Jon, Ben Jacobs, and Paul Lewis. "Baltimore Protests Turn into Riots as Mayor Declares State of Emergency." *Guardian*, April 28, 2015. <https://www.theguardian.com/us-news/2015/apr/27/baltimore-police-protesters-violence-freddie-gray>.
- Terwindt, Carolijn. *When Protest Becomes Crime: Politics and Law in Liberal Democracies*. London: Pluto Press, 2020.
- Thompson, Christie. "Hacktivism: Civil Disobedience or Cyber Crime?" *ProPublica*, January 18, 2013. <https://www.propublica.org/article/hacktivism-civil-disobedience-or-cyber-crime>.
- Waddington, David. "The Madness of the Mob? Explaining the 'Irrationality' and Destructiveness of Crowd Violence." *Sociology Compass* 2, no. 2 (March 2008): 675–87.
- Walzer, Michael. *Political Action: A Practical Guide to Movement Politics*. Chicago: Quadrangle Books, 1971.
- Young, Iris Marion. "Activist Challenges to Deliberative Democracy." *Political Theory* 29, no. 5 (2001): 670–90.

## SCOPE RESTRICTIONS, NATIONAL PARTIALITY, AND WAR

Jeremy Davis

**M**OST PEOPLE believe partiality is justified among friends, parents and children, family members, and romantic partners. These are the paradigmatic cases of partiality. But many of us also think partiality can be justified in a broader range of cases as well—for example, toward those with whom we engage in certain shared projects, our colleagues, people in our local community, our co-unionists, and our conationals, among others. On this commonsense approach, partiality applies within a wide range of relationships—not just the more intimate relationships that characterize the paradigmatic cases but also certain relationships that consist of membership in a certain sort of collective.<sup>1</sup>

Given this broad application, it is no surprise that there are important differences in the way partiality applies in these various relationships. For one thing, not all relationships give rise to reasons of partiality of the same strength. The extent to which such reasons can outweigh other competing reasons—in particular, one's impartial moral reasons—will vary according to certain features of the relationship in question. Another difference, and the one on which I will focus in this essay, concerns the scope of the application of partiality within certain relationships. In particular, I will argue that certain relationships give rise to reasons of partiality that are *scope restricted*. To say that a given relationship generates scope-restricted reasons just means that it gives rise to such reasons only with respect to certain goods or in certain contexts. Recognizing the scope-restricted nature of certain forms of partiality is important for understanding the structure of the phenomenon of partiality, as well as its application in a wide range of contexts.

But it also serves as the basis for a rebuttal to a popular objection to the appli-

<sup>1</sup> This is not to say that all of these relationships will give rise to partiality in every case. For one thing, many of these relationships will be purely instrumental, morally toxic, or otherwise devoid of significance to the parties within them. When this is true, partiality may indeed be unjustified. The point is simply that these relationships sometimes give rise to reasons to be partial.

cation of national partiality in war. According to the objection, if national partiality is justified in war, then it must be the case that we can prefer our conationals in ordinary interpersonal cases as well—for example, in forced-choice scenarios, like trolley problems. And yet, national partiality does not seem justified in many of these cases. Therefore, national partiality cannot justify prioritizing conationals in war. I argue that this analogous reasoning is misguided: the paradigmatic relationships are generally scope unrestricted, while national partiality is scope restricted. The two kinds of partiality are therefore not straightforwardly analogous. Indeed, we should expect that there are many cases in which national partiality is unjustified—namely, the cases that fall outside the relevant scope of that relationship. When a case falls within the scope, however, partiality may indeed be justified. I argue that war is one central case that falls within the scope of national partiality.

In the next section, I explain the conditions under which certain relationships give rise to partiality that is scope restricted. Then, in section 2, I show that partiality among conationals is another example of scope-restricted partiality. Finally, in section 3, I take up the aforementioned objection and show how our understanding of national partiality as a scope-restricted form of partiality serves to overcome that objection.

### 1. SCOPE RESTRICTIONS

To understand how reasons of partiality might be scope restricted, it is helpful to look at a specific example. Take the colleague relationship: I have reasons of partiality toward my colleague, which extend to the goods and the context relevant to the particular relationship we share, such as those relating to our productive lives and our particular industry. Indeed, these reasons might be quite strong in some cases: they might give me reason to break otherwise significant promises to others, to divert financial resources to my colleague instead of others, and so on. And yet, these reasons of partiality arise only with respect to certain interests and within the context of our relationship as colleagues. Put differently, it is not as though all of my colleague's interests have special salience for me. I do not have reason to, for example, break important promises to others or divert financial resources to him instead of others so as to promote his athletic or domestic interests, since neither of these is a good or context relevant to the particular relationship we share.<sup>2</sup>

2 Some may wish to distinguish between *associative duties*, which are a type of agent-relative duty based in a special relationship, and *special duties*, which are based in certain kinds of interactions (e.g., promises or creation of expectations). And perhaps this distinction helps us

Moreover, even those reasons that do fall under this purview might cease to apply outside the context of the relationship. For example, even though I have reasons of partiality to promote my colleague's work-related interests, these only apply within the context of our shared work life. I do not have reason to promote these interests of his when, for example, he takes another job in the same field, or even when he takes a second job doing similar work. In many cases, the context is limited to interactions via particular institutions, of varying degrees of formality, such as a government or other collective body; a system of laws, customs, or practices; and so forth. Those in relationships of this sort have reasons of partiality to act through these institutions that do not apply in the absence of such an institution (whether in general or in a particular situation). For example, the co-unionist relationship is governed in part by shared membership in the union. Co-unionists therefore have reasons of partiality toward one another with respect to a narrow set of interests only when acting through or within that institutional arrangement—for example, to promote co-unionists' productive and economic life in that particular domain. All else being equal, co-unionists do not have such reasons when the institution is not the means through which the goods would be promoted. So, while I may have reason to promote the economic life of my fellow unionists in the context of collective bargaining as members of the union, I do not have a similar reason to promote the economic life of my fellow unionists by helping them with consolidating their debts, refinancing their mortgages, or making investments. Though such issues do indeed constitute part of their economic life, they are not part of the institutional arrangement that forms the core of our relationship. These interests, therefore, fall beyond the scope of our relationship.

So, the scope of one's reasons of partiality is restricted both to a certain set of interests and to a particular context of interaction. The particular type of relationship in question delimits the interests that have special salience; some interests will have special salience in some relationships and not others. And these interests only have special salience within the context of the relationship; outside this context, these interests ought to be considered only impartially.

Several other philosophers have recognized this feature of certain special re-

---

to see why some relationships will turn out to be scope restricted and others not: the former are cases of special duties, while the latter are cases of associative duties. If this is true, then much of what I say about scope restrictions in fact just tracks the different kinds of duties in question. I do not have the space here to treat this worry completely. It will suffice to note that I think all the duties I discuss here are of a similar sort; they are all a kind of associative duty. I do not think the colleague or co-unionist relationship can be understood entirely as a mere promise or creation of expectation; I think there is something more morally important about such relationships.



relationships. In his discussion of associative duties, Ronald Dworkin says, “my concern for my union ‘brother’ is general across the economic and productive life we share but does not extend to his success in social life, as my concern for my biological brother does.”<sup>3</sup> Sarah Stroud makes a related point:

If the cellist in my quintet needs a new bow and is too poor to buy one, I might be permitted to direct money that could otherwise go to famine relief to the cause of getting her one, in so far as that makes possible the continuation of a collective project in which I am engaged: playing the Schumann Quintet. But if she needs money for reasons unrelated to the quintet project, then it seems I can send my money to famine relief without any cost to my (or our) projects; so ... I would not have the same moral case for directing it to her instead.<sup>4</sup>

These examples nicely illustrate the restriction on interests. As Dworkin’s example shows, the salient interests of one’s co-unionist are only those that are central to the relationship—namely, those that are part of the “economic and productive life [they] share.” Many of his other interests (e.g., his social life) are simply not salient to their relationship and therefore fail to generate reasons of partiality with respect to them. Stroud’s example proceeds in a similar fashion. The only interests of my quintet partner that have special salience for me are those related to our relationship—or as Stroud would put it, those that form part of the particular joint project we share. Her interests that do not concern this project are not eligible for partiality.

These examples can also be adapted to highlight the importance of context. Consider the case of the cellist. While her musical needs have special salience to me within the context of our relationship, these same interests lose such importance when they are outside the relationship. For example, there are certain musical needs she may have such that though they would generate reasons of partiality were they to be relevant to our quintet, they fail to do so when they are for some entirely unrelated project.

By contrast, certain other relationships do not seem to give rise to reasons that are restricted in this way; call such reasons *scope unrestricted*. The clearest case of this is the reasons parents have toward their children. Intuitively, a father’s reasons of partiality toward his son do not concern only matters relating to developing and preserving their special relationship: he also has reasons that extend to his son’s health and general welfare, life prospects, happiness, education, and so on. The same seems true of intimate romantic partners and friends.

3 Dworkin, *Law’s Empire*, 200.

4 Stroud, “Permissible Partiality, Projects, and Plural Agency,” 148.

While I certainly have strong reasons to attend to those matters uniquely concerning our romantic partnership or friendship—our mutual treatment of one another, our ongoing promises and agreements, and so on—my reasons of partiality also extend to other matters of my partner or friend’s personal life, including her general welfare, her future, and so on.

One might deny that certain relationships are scope unrestricted. Even in the closest relationships, one might argue, the interests that can give rise to reasons of partiality are indeed limited in significant ways. For certain interests—perhaps those concerning the person’s other relationships, private life, or other intimate affairs—one might rightly say, to put it colloquially, “that is none of my business.” There are several different interpretations of this point. One might claim that some interests will fall outside the scope of even the paradigmatic relationships. But I doubt we will find a case of an interest of one’s child that is, as a general matter, *never* the business of the parent at all—and likewise for the other paradigmatic relationships. Surely, we ought to respect our loved ones’ privacy, avoid meddling too much in our friends’ love lives, and give our children the space to make their own choices without our interference. But it does not follow from this that these interests are beyond the scope of our relationship. All this shows is that certain relationship goods prohibit certain forms of intervention.<sup>5</sup>

One might instead claim that while certain interests have special salience to us, and give rise to reasons of partiality, they are sometimes *silenced* by the presence of other important duties in particular contexts. For example, a police officer has (*qua* parent) reasons of partiality toward her son with respect to his well-being, but these reasons are silenced when determining whether she ought to arrest him for a serious crime of which he is rightly accused.<sup>6</sup> Surely there are other cases that take this general form. If so, then reasons of partiality might appear to be restricted in scope in certain cases due to the competing duties inherent in one’s other social roles.

I am unsure whether reasons of partiality in such cases are in fact silenced, rather than significantly outweighed, by competing duties.<sup>7</sup> But I will grant the

- 5 Indeed, the duty of privacy is, in an important sense, another reason of partiality: surely we have a stronger *pro tanto* duty to respect the privacy of those with whom we share a special relationship, especially one based on trust. Thus, this noninterventionist approach is not just compatible with but also constitutive of the relationship.
- 6 One might be inclined to view certain professions or positions as roles, which come with their own distinctive moral obligations. For a prominent defense of role obligations, see Hardimon, “Role Obligations.”
- 7 It could be that while the officer has a duty to enforce the law fairly, she is nevertheless permitted to give some special treatment to her son with respect to minor violations (e.g., speeding tickets), particularly when the cost to his well-being is significant. If so, this would

point here as it applies to certain social roles, such as police officer, judge, elected official, and others. Still, this point does not significantly threaten the claim that certain relationships are scope unrestricted. First, the claim about silencing does not seem true of ordinary relationships absent any such competing duties from social roles: while parents generally ought to avoid interfering in their children's lives, when the stakes are higher their intervention seems not just permissible but also required. If such reasons were silenced, this would not be the case. This suffices to show that at least *some* relationships are scope unrestricted. Furthermore, it is not the case that the police officer's relationship with her son is inherently scope restricted; were she to retire from the force tomorrow, no such restrictions would remain. It is, therefore, the particular combination of her competing duties—to her son, on the one hand, and to the code of her profession, on the other—coupled with the apparent supremacy of the latter over the former that explains the moral restrictions she faces.

It is worth noting that the foregoing points concern only the general structure of partiality. If it were to turn out that all relationships are scope restricted after all, the argument in the following sections would still go through. What would be important in this case would be the extent of the scope restriction with respect to these paradigmatic relationships. And it should be clear enough, given what I have said so far, that if these relationships have any such restrictions at all, they are much less extensive than those in the examples of scope-restricted relationships I mentioned above—in particular, national partiality.

It is worth noting two further aspects of scope-restricted partiality. First, it does not follow from the fact that partiality in a given case is scope restricted that such reasons are therefore weaker than in scope-unrestricted cases. Of course, they may sometimes be weaker: my reasons of partiality to the cellist in my quintet are generally much weaker than, and would often be outweighed by, my reasons of partiality toward my romantic partner. However, this may not always be true. Indeed, my reasons to confer special treatment on, for example, my union brothers may, in certain contexts, outweigh some of my other reasons of partiality (e.g., to my friends or family), and they may even override certain demands of morality more generally—perhaps even significant demands. Thus, a given reason's weight cannot be entirely determined by its scope.

A second important point is that determining the exact scope for a given relationship will be difficult, given that the boundaries in some cases are quite vague.

---

suggest that reasons of partiality to her son are not silenced, but rather outweighed, by the significance of the opposing duties inherent to her role. I will not pursue this point further here, since it requires a more thorough discussion of the particular features of such roles, and does not bear too significantly on the central argument.

Often this is because it may be unclear which specific interests fall within the scope for a given relationship. Does the fact that the co-unionist relationship involves shared membership in a collective concerned with members' economic and occupational lives give rise to reasons of partiality for one co-unionist to help another with tending to her work-related injury or learning new occupational skills, since such things could plausibly be thought to extend to her occupational life? While one's personal investments are generally outside the scope of the co-unionist relationship, what about when these investments are importantly tied to one's role as a laborer? Moreover, reasonable people might disagree about the scope in a given case. For example, David Miller claims that one's "collegial obligations extend to general human interests, so that if there are two students who need to be driven urgently to the hospital, and I can only take one, then ... I ought to give priority to the one who belongs to my college."<sup>8</sup> This claim appears to rest on his belief that the collegial relationships within certain colleges (i.e., his own) have a broader scope, perhaps (though this part is left unexplained) by virtue of their more tightly knit community.<sup>9</sup> This strikes me, and I suspect many others, as implausible. But we need not resolve it here; while answers to these and other similar questions are necessary for rendering a precise verdict in specific cases, they are not required for the purposes of the present discussion, which seeks only to show that such scope restrictions do exist and can be explained by these general features of the relationships.

It may also be difficult to determine whether one has reasons of partiality in a given scenario because certain relationships (e.g., those among colleagues, which give rise to reasons of partiality that are scope restricted) are often also inchoate friendships, and friendship is a paradigmatically scope-unrestricted relationship. Consider once more the co-unionist from my adapted version of Dworkin's example. As I said, *qua* co-unionist, I do not have reasons of partiality with respect to his economic affairs outside of the context of our co-unionist relationship; *qua* friend, however, I may indeed have such reasons—though perhaps such reasons in cases like these are quite weak in general. Thus, as scope-restricted relationships evolve in the direction of scope-unrestricted relationships, this will give rise to an expansion of the scope of the reasons of partiality that applies to them.

How do we know the scope of reasons for scope-restricted relationships? The most natural thought is that what grounds partiality can also be applied as its

8 Miller, *On Nationality*, 66.

9 This is perhaps based on the idea that at some colleges, professors and others in supervisory roles are required to act *in loco parentis* with respect to the students in their college. If true, this would explain the extent of the partiality in this case, though I am skeptical that it applies much beyond that. Thanks to Arthur Ripstein for suggesting this point.

scope. Consider again Stroud's view, which grounds reasons of partiality in joint agency: acting together on a shared project gives individuals reasons to confer special treatment on those with whom they pursue that project.<sup>10</sup> On Stroud's account, scope restrictions are effectively built into the grounds: this special treatment that individuals may confer on one another extends only within the parameters dictated by that joint project. Reasons of partiality cease to apply outside the context of the project that gave rise to them.<sup>11</sup>

Another popular view holds that partiality is grounded in certain relationship goods.<sup>12</sup> One of the most prominent defenders of this view is Jonathan Seglow, who argues that "participants in social relationships uniquely enjoy certain relationship goods, and their associative duties involve promoting those goods."<sup>13</sup> These goods vary according to the relationship, but the ones most relevant for our purposes are what Seglow calls "common purpose goods," which are "embedded within the purposes of those associations which seek to express and promote them."<sup>14</sup> Like Stroud's account, Seglow's account also draws a clear connection between partiality's grounds and its scope: the very same goods that ground our reasons of partiality also serve to limit the extent of the application of those reasons.

Finally, many philosophers ground partiality in a certain sort of shared history. C. D. Broad holds that this special consideration is justified by virtue of "the traces of innumerable actions and experiences in common."<sup>15</sup> Thomas Hurka argues that partiality is justified on the basis of a shared history of doing good (or suffering evil), and the extent of this justification is determined by the degree of interaction and the good produced in the relationship: the greater the degree of interaction and the greater the good produced, the stronger the reasons to

10 See also Gilbert, *Living Together*.

11 Stroud, "Permissible Partiality, Projects, and Plural Agency," 148. Ultimately, I think this account is too narrow, particularly in terms of context. It seems to rule out cases in which one could benefit another in an area that falls just outside their joint project. For example, the collective project in the example is defined strictly as "playing the Schumann Quintet," and she no doubt has reasons of partiality with respect to that shared project. However, it seems intuitively plausible that she also has reasons of partiality that extend just beyond that project—say, to help her work on a section of another piece she will soon perform as part of another ensemble. This does not bear centrally on the argument that follows, so I will set it aside.

12 Keller, "Four Theories of Filial Duty"; Seglow, *Defending Associative Duties*; Swift and Brighthouse, *Family Values*.

13 Seglow, *Defending Associative Duties*, 2.

14 Seglow, *Defending Associative Duties*, 119.

15 Broad, *Examination of McTaggart's Philosophy*, 2:138.

be partial.<sup>16</sup> And Niko Kolodny argues that our reasons of partiality resonate with the good of the shared history of encounter.<sup>17</sup> While each of these versions differs slightly from the others, the central point on which they all agree is that reasons of partiality are grounded in a good shared history.

The shared history of a given relationship can also serve as its scope. The history I share with my co-unionists is limited to the economic and social roles we inhabit as members of a trade union. That is, our shared history is limited to a particular set of interests in a particular context. In ordinary cases, my history with a given co-unionist does not extend to her family life, her moral development, or other interests unrelated to our shared vocation. Nor does our shared history extend to certain otherwise relevant interests—namely, those pertaining to her economic life—when they arise in an unrelated context (say, her inheritance wealth, home value, or other sources of income). The same point applies to the cellist in my quintet: our shared history concerns our mutual interests in playing music and developing our musical talents and so can plausibly extend to related contexts, such as performing in other groups.<sup>18</sup> But it does not extend to other interests that are not part of that history, such as those concerning her marital relationship, spiritual life, or athletic endeavors.

The projects view and the relationship-goods view both yield a scope of reasons that involves promoting the very goods that ground partiality.<sup>19</sup> The shared-history account does not do this: it is not that the cellist has reason to

16 Hurka, “The Justification of National Partiality.”

17 Kolodny, “Which Relationships Justify Partiality?”

18 Indeed, given the broad nature of the good in question, I suspect we would think it perfectly appropriate to devote special attention to helping the cellist (when possible) with a large range of musical interests—for example, practicing her drum rudiments, working on vocal melodies for her pop duet performance, or picking out a good used synthesizer—even if this would involve forgoing opportunities to help others, perhaps even many more, in similar ways.

19 One might wonder whether the fact that these relationship goods are dependent on certain underlying goods means those latter goods must also qualify as falling within the scope of the relationship. For example, the cellist in my quintet can only perform well provided that her mental health is cared for, her economic resources are adequate, and so forth. I admit that cases of this sort highlight an ongoing challenge for my view concerning how to delineate precisely what falls within the scope of a given relationship. One plausible, though imperfect, response is to hold that certain underlying elements—such as one’s mental health—are in most cases mere preconditions for the relationship and not a part of the relationship *per se*. This would explain why mental health more broadly is beyond the scope, but a specific case of anxiety-driven yips might fall within the scope. Moreover, in many cases, even if the goods are indeed relevant, the context of the interaction will explain why they lie beyond the scope of the relationship. (Thank you to an anonymous reviewer for suggesting this point.)

*promote* her shared history of interaction with the members of her quintet. Rather, on certain versions of the shared-history account, the goods one has special reason to promote are those that resonate with the goodness of the history, and the context of the shared history provides the context of the relationship in which those goods have particular salience.<sup>20</sup> Paradigmatically scope-unrestricted relationships, like long-standing friendships or romantic relationships, will have such broad and extensive shared histories that the context is effectively unlimited.

There are still other accounts of partiality that could extend in this way to delimit the scope of reasons, but these three will suffice for our purposes.<sup>21</sup> I remain agnostic here about which of these is the best account of partiality's grounds. The arguments to follow are broadly compatible with all three accounts just discussed.

## 2. NATIONAL PARTIALITY

If the foregoing arguments are correct, then some relationships generate reasons of partiality that are scope restricted, while others are scope unrestricted. What should we say about national partiality? We can understand the nation as a political community centered on some cluster of the following: shared institutions, societal norms, cultural traditions and values, language, laws, and shared recognition of their relationships among themselves. This remains imprecise, of course, but we need not settle on any narrow view of what the nation consists in for our present purposes; our commonsense understanding of this idea should suffice for now. Conationals, on this view, are just people who share membership in the nation, however we understand this idea.

Are the reasons of partiality among conationals more like those that exist between parents and children and between close friends, in that they extend to virtually all matters in that person's life, or are they more like those that exist among members of a union or members of an orchestra, in that they extend only to those matters directly related to their comembership?

In my view, the conational relationship has much more in common with the relationship between co-unionists than it does with the relationship between parents and children. Just like co-unionists, conationals share many important political projects and values within the context of which partiality is often appropriate. Of course, the scope of the conational relationship is generally much wider

20 Hurka, "Love and Reasons"; Kolodny, "Which Relationships Justify Partiality?"

21 There are also pluralist views, according to which there are several grounds of partiality. For one recent example of a view like this, see Lord, "Justifying Partiality."

than that of the co-unionist relationship: while the co-unionist relationship only extends across the economic and productive life the two share (as Dworkin puts it), the conational relationship extends beyond these to include a broader range of interests—namely, those concerning individual welfare, health, general safety, education, and so on. The precise extent of the relationship will be determined and constrained in each case by the certain features of the nation in question. To see this, we can look to the history that conationals share or consider what constitutes their relationship in the first place. This relationship centers on a shared history of creating and sustaining important institutions; preserving and protecting a certain way of life, which for many is a key component of their identity; and focusing efforts on securing and protecting individuals' core rights and interests.

And yet, despite this rather broad range of interests that are central to the conational relationship, partiality among conationals seems restricted in a way that partiality toward one's child (or dear friend or spouse) does not. In particular, though the broad range of interests of one's conationals gives rise to reasons within the context of their shared life, such as the institutions they share, these interests cease to give rise to such reasons outside such contexts. To see this, consider the following scenario:

*Desert Island:* You and I are conationals who happen to be traveling on the same flight across the South Pacific, each of us taking completely independent vacations. Our plane crashes on a remote island, and you and I now find ourselves among the dozens of survivors of many different nationalities. As we scrounge to survive, each of us finds ourselves in a position (though with no particular authority) of being able to divert certain scarce resources, which are necessary for survival, to some but not all of the victims. (Assume that everyone's needs with respect to these resources are roughly the same.)

Does the mere fact that we share a nationality now give me special reason to allocate scarce resources to you over the others? (Does the mere fact that certain other survivors also share a nationality give them such reasons to favor each other?) Am I permitted to help rescue you over any of the others? Many will say no. One view, of course, is that national partiality is unjustified in general when it comes to cases of this sort, no matter the particular story. This is roughly the cosmopolitan approach that we will see later on in section 3.

But another response to this case allows for a general endorsement of national partiality and yet preserves the intuition that conationals are not morally permitted to prioritize one another *qua* conationals in this case. This response appeals to the scope-restricted nature of the relationship and, in particular, the



importance of the context of interaction. While the goods of survival and care give rise to reasons of partiality between us within certain contexts, such as those falling under the remit of our shared institutions, the current context is importantly different: the fact that we are conationals is completely incidental to our current circumstances. It is not the restriction on the goods, then, but the restriction on the context that limits the scope of conational partiality in cases like this.

Moreover, none of the underlying moral justifications listed above—joint projects, relationship goods, or a shared history—seem to apply here. Our vacations have nothing to do with our joint projects, nor do they provide a context in which we realize our distinctive conational relationship goods, and the context does not fall within our shared history as conationals. Put differently, our conationality is not morally salient for partiality in this scenario: I no longer relate to you *qua* conational, but rather *qua* individual—that is, just as I relate to the other survivors of the accident. This conclusion applies similarly to other scope-restricted relationships—for example, if instead of being conationals stranded on a desert island, we were co-unionists, or played in a quintet together.<sup>22</sup>

But we would draw an entirely different conclusion if the relationship in question were scope unrestricted.<sup>23</sup> Suppose that instead of being conationals, you and I are siblings, close friends, romantic partners, or father and son. In any of these cases, our judgment would surely shift: it seems clear that I *do* have special reason to allot scarce resources to you and to prefer rescuing you over others. This is because these interests of yours are not restricted to some specific context. It is not the case that we merely happen to encounter each other here as siblings or friends, as in the conational case.<sup>24</sup> Indeed, the central feature of scope-unrestricted relationships is that they do not seem incidental in this way.<sup>25</sup> This idea is supported by the fact that the underlying moral justification of the relationship is essentially unrestricted: the joint projects, relationship goods, and

22 For some of these cases, it will not be the different context itself that rules out partiality, because the interest in question does not have special salience in any context. I think it is plausible that this is true of the members of a quintet, but we need not settle this point here.

23 I assume here that my earlier argument concerning the possibility of scope-unrestricted relationships was sufficiently persuasive. If not, then this point should be amended to reflect the idea that such relationships are indeed scope restricted, though such restrictions are many fewer in number. The point I am making here is not ultimately affected in any significant way by this change.

24 Notice that this is true even if we modify the example so that we were not traveling together, and only discovered we were on the same flight after it crashed.

25 Though, of course, one can imagine fanciful cases in which one discovers that another passenger is one's biological brother. This mere biological relationship is not what I (or, I suspect, anyone) has in mind when they imagine the moral relevance of the sibling relationship.

shared history of a genuine friendship are not limited to a particular context. The pervasiveness of their application is just part of what a genuine friendship is.<sup>26</sup> But again, we would not say the same for the conational relationship.

As I noted, one important feature of *Desert Island* is that it is an example of conational interaction in the absence of, among other things, any of the institutions that form part of the broader conational relationship. However, if we change the example such that the conational interaction in question is mediated through one of the conationals' shared institutions, we get a different verdict.

*Desert Island Rescue:* Your plane crashes on a remote island, and you now find yourself among the dozens of survivors of many different nationalities, some of whom are your conationals. There happens to be a small naval vessel in the vicinity composed of members of your nation's navy. Upon learning of the crash, this naval ship comes to rescue you and your other conationals. Unfortunately, they only have room and supplies for a limited number of additional passengers. They can take you and your other conational survivors, or else a group of other survivors chosen at random.<sup>27</sup>

I suspect many would accept that this crew has reasons of partiality to rescue their conationals, rather than a random group of survivors. But now suppose that instead of being a naval vessel, it was a random fishing vessel that happened to be composed of members of your nation. I think many would think the crew aboard this ship does not have a compelling moral reason to rescue only their conationals. Morality would require that they rescue based on the morally relevant factors—for example, who is in most need of care, regardless of nationality.

One way of explaining this pair of judgments is that the navy, unlike the fishing crew, is part of one of your nation's shared institutions, and interactions between conationals via institutions they share take on a different character from interactions outside of those institutions. That is, these shared institutions can make a context that would otherwise be outside the scope of the relationships that fall within it. This is not just limited to cases involving the military. Indeed, the same point applies with respect to certain other institutions we share, such as a health-care system: we have strong reasons to care for our conationals'

26 I use the term "genuine" here to distinguish from more casual cases—for example, the "friends" you happen to play tennis with at the tennis club or your inchoate friendships where calling someone a "friend" is partly intended to be proleptic.

27 Rescues of this sort are not entirely uncommon. For one recent example, see the evacuation of nearly fifteen thousand American citizens in Lebanon at the onset of the Lebanese civil war in 2006. The details of this case differ slightly from those in *Desert Island Rescue*, but the cases are sufficiently similar in all the most morally salient ways.

health via the institution we share that serves this purpose, but we have no such reasons of partiality outside of that context.<sup>28</sup> It might be thought that the members of the institution in question—in our above example, the members of the navy—are, in an important sense, an extension of the nation itself. As such, this example might look like a top-down version of partiality rather than partiality among conationals. But notice that even other conationals may have reasons of partiality of a relevant sort. That is, I have reasons of partiality to promote and fund (e.g., with my tax dollars) the ongoing efforts of the navy with respect to its mission to help my conationals abroad. In other words, I have reasons of partiality to promote the existence and efforts of the shared institution. (The same can be said for other forms of international protection, such as embassies.) I do not have such reasons to promote similar efforts by fishing crews who head out into international waters. Thus, while it is true that, in this case, the members of the navy encounter their conationals in part as an arm of the nation, this does not exhaust the ways of understanding the role that shared institutions play in shaping the context of conational interaction.

One might object that those on the fishing vessel also have reasons of partiality to rescue you and your conationals. After all, they have benefited from your nation's economic and educational institutions, perhaps even in relation to their particular expedition. Do not those facts provide sufficient context for our current interaction?<sup>29</sup> It is important to remember, however, that the mere existence of shared institutions does not determine whether a given interaction falls within the scope of the relationship; this would make the issue of context functionally irrelevant, since all interactions between conationals would inevitably qualify. Nor is it determined simply by the fact that there is some causal connection between these institutions and our current context. This, too, would be overbroad for the very same reasons: the causal connection between overseas visits like those of the fishermen and, for example, state-provided elementary education, government administration, and interstate highways is far too weak to serve as grounds for a relevant context of interaction between conationals.

The previous example is one in which the actions of insiders to the relationship cause the context to shift, thereby altering the reasons of partiality. But actions of outsiders to the relationship can cause this shift as well. To see this, consider the following example:

*Terrorist Vacation:* There are two neighboring, relatively isolated beach re-

28 This is not to say that we would not also have reasons to care for their health if the institution in question were solely lacking or if it were to fail in its efforts in certain ways.

29 Thank you to an anonymous reviewer for raising this objection.

sorts: one happens to be populated entirely by vacationers from Nation A; the other happens to be populated entirely by vacationers from Nation B. By coincidence, two distinct terrorist groups attack the two resorts: Group 1 attacks the resort populated by Nation A; Group 2 attacks the resort populated by Nation B. Furthermore, suppose that the terrorist groups chose their respective targets because of the specific populations at those resorts: Group 1 wanted to attack members of Nation A; Group 2 wanted to attack members of Nation B. As with Desert Island, both groups are in a position to divert scarce resources, attention, and aid to some but not all of those in both victim groups who have been injured by the attacks.

While the example is (admittedly) fanciful and unrealistic, the basic question it raises is important: Does the fact that one's nationality was central to the causal story explaining one's predicament change the moral salience of the conational relationship in that setting?<sup>30</sup> I think it does: the fact that members of Nation A were attacked *because* they are members of Nation A creates a context for their conationality where there would not have been one in the absence of that aim.<sup>31</sup> Their relationship becomes salient because others have chosen it to be salient by virtue of the particular aims behind their actions. As a result of this, the victims of the attack have reasons of partiality toward their conationals, while they would not have had such reasons had they been targeted for some unrelated reason (e.g., mere opportunism). This particular example is unlikely to arise in this pure form, but the feature it isolates and highlights is much more pervasive. In the next section, I argue that war provides a similar context in which the conational relationship can become salient by virtue of the actions of others.

### 3. SCOPE RESTRICTIONS AND WAR

The foregoing discussion has shown that scope restrictions are an important structural feature of partiality in general and national partiality in particular. But

30 Though the specifics are quite different, it is a commonly accepted idea that reasons of partiality—at least of a sort—might arise from the fact that others have made your group identity relevant. For a discussion of solidarity in the Black community, which bears some similarities to this point, see Shelby, *We Who Are Dark*.

31 Notice that our judgments might shift slightly if the group's composition were more mixed—i.e., with some of Nation A and some of some other nation. In this case, the context of the conational relationship is active, as it is in the purer case, but we might have reasons of a different sort to tend to those who were caught in the middle of an attack against us. It might therefore be that, all things considered, we do not have greater reason to attend to the needs of our conationals over the other victims.

they are more than just a relatively overlooked structural feature of partiality. Indeed, as I suggested at the outset of this essay, this idea also provides some of the resources for avoiding one of the most common objections leveled against the application of national partiality in war.

Several proponents of national partiality believe that nations may confer greater weight on the lives of their own citizens than on the citizens of the enemy state in the context of war.<sup>32</sup> One strategy for defending this claim proceeds by analogy from other cases of justified partiality, such as the relationship between parent and child. For example, Thomas Hurka considers a case in which a victim is attacked by an aggressor, and a third party can rescue the victim only by throwing a grenade that will kill, as an unavoidable side effect, one innocent bystander. If the third party is not in any special relationship with the victim, then he is not permitted to throw the grenade. This changes, however, if the third party is the victim's father. In this case, Hurka says, "it seems to me that he may throw the grenade, and may do so even if this will kill some number of bystanders greater than one. If he is not aiming at the bystanders but killing them collaterally, he may show some preference for his daughter." Hurka then goes on to claim that this case is analogous to the partiality that conationals may show in the context of war: "when weighing its own civilians' lives against those of enemy civilians it will merely collaterally kill, a nation may give some preference to the former."<sup>33</sup>

Some philosophers have objected to Hurka's reasoning. David Lefkowitz argues that the right approach to determining whether national partiality is justified in war is not to reason by analogy from cases of parent-child partiality to wartime cases of partiality among conationals, but rather to consider whether national partiality is justified in certain nonwar contexts. In his view, the proper nonwar scenario for testing the strength of national partiality is a case in which one can save a conational at the cost of one non-conational bystander collateral death: "As for Hurka's non-war scenario, a variation on it more closely analogous to choosing between harm to compatriot or to enemy non-combatants would involve throwing a grenade that will kill one or even several foreigners in order to save one compatriot. I contend that such an act is not morally justifiable."<sup>34</sup> In other words, the mere fact that certain individuals share a nationality is insufficient to outweigh the competing moral demands to others in this kind of inter-

32 Bazargan-Forward, "Weighing Lives in War"; Hurka, "Proportionality in the Morality of War"; Kamm, "Failures of Just War Theory"; Lazar, "Associative Duties and the Ethics of Killing in War"; McMahan and McKim, "The Just War and the Gulf War"; Miller, "Reasonable Partiality towards Compatriots."

33 Hurka, "Proportionality in the Morality of War," 61.

34 Lefkowitz, "Partiality and Weighing Harm to Noncombatants," 307.

personal case, even if such duties might be outweighed when the relationship is, say, between parent and child.<sup>35</sup> Therefore, the conational relationship cannot outweigh those same sorts of demands in war cases.

Cécile Fabre has raised a similar objection. She argues that what she calls “patriotic partiality” does not generate special permissions in interpersonal cases, and so we can infer that it has no role to play in wartime cases either.<sup>36</sup> To support her claim, she employs the following example:

Suppose that André, in the course of defending his life, has a choice between killing Carl and foreseeably killing Bernard, and killing Carl and foreseeably killing Werner, a German bystander who is wholly innocent of Carl’s wrongdoing. According to cosmopolitan justice he may not confer greater weight on Bernard’s life than on Werner’s simply on the basis that the former is, whilst the latter is not, a compatriot.<sup>37</sup>

David Miller has argued in a similar vein against national partiality’s ability to justify violating certain duties that are characteristic of war. He claims that while national partiality does give rise to certain permissions to favor one’s conationals, national partiality cannot justify violating serious negative duties toward others, like the duty to avoid infringing their basic rights. Referencing the trolley problem, made famous by Philippa Foot and Judith Thomson, Miller writes: “I don’t think it would be justifiable to switch the trolley from a track on which it was hurtling towards a compatriot on to a track on which it would hurtle towards a foreigner. . . . At this level, morality appears to me to require strict impartiality at least as far as nationality is concerned.”<sup>38</sup> Miller’s reasoning is just like that of Lefkowitz and Fabre: we can tell whether national partiality can justify infringing certain duties that are characteristic of war—for example, killing or severely harming others—by appealing to interpersonal cases among conationals, like trolley cases. While Miller does not refer to war specifically in

35 To be clear, Lefkowitz does not claim specifically that throwing the grenade is justified in these interpersonal cases. My point is simply that one could, and many likely do, hold this pair of judgments consistently.

36 Fabre does, however, grant that a certain kind of “patriotic partiality” is permitted in war, but it does not involve the sort of partiality with which we have been heretofore concerned (i.e., attributing special moral status to someone in virtue of shared membership in a nation). Rather, it involves defending jointly held rights, and thus makes no necessary reference to the moral importance of one’s conational relationship.

37 Fabre, *Cosmopolitan War*, 85.

38 Miller, “Reasonable Partiality towards Compatriots,” 74–75. For the two most classic discussions of the trolley problem, see Foot, “The Problem of Abortion and the Doctrine of Double Effect”; and Thomson, “The Trolley Problem.”

this discussion, his claim picks out the same general category of duties, which are central to war, as Lefkowitz and Fabre pick out, and his claim clearly extends to the domain of war.

The view shared among these philosophers is essentially that national partiality in the context of war is impermissible if the same sort of national partiality is not permissible in a relevantly similar interpersonal context. Call this the *interpersonal cases argument*. If this argument is correct, Hurka's view and others like it proceed by way of a faulty analogy to a mistaken conclusion.

Some who are sympathetic to Hurka's view may wish simply to reject the intuitions pumped by Lefkowitz's, Fabre's, and Miller's examples. But let us grant here that it is wrong to prefer our conationals over foreigners in certain interpersonal contexts, like those in the above examples, even though this is not the case with certain other relationships, like the parent-child relationship. (If this is correct, then we must concede the first part of Lefkowitz's argument: *pace* Hurka, national partiality is not so closely analogous to parent-child partiality after all.)

And yet, in granting the claim that conationals are not permitted to be partial to one another in the nonwar scenarios described above, one is not thereby committed to the interpersonal cases argument's main conclusion—namely, that national partiality is impermissible in war. We cannot draw conclusions about the justification of national partiality in some cases from the (im)permissibility of national partiality in other contexts. This is because, as I argued above, national partiality is scope restricted: the conational relationship gives rise to reasons only in particular contexts.

To prove that the interpersonal cases argument is false, however, it is not enough to show that national partiality is scope restricted. That national partiality is scope restricted only shows that it *does not* follow that national partiality *is not* permitted in war. To show that it *is* permitted in war, as Hurka has argued, we must also defend the claim that war is one of the contexts in which the conational relationship is sufficiently morally salient.

Consider what the various possible grounds of partiality discussed earlier suggest about the context of war. War typically involves threats to conationals' joint projects of national defense against unjust aggression; the relevant relationship goods, including self-determination, protection from outside aggressors, and the preservation of a collective democratic life; and their shared history of promoting and developing their shared life across generations. All of these goods involve collective deliberation about group and individual political rights and values, and they take place within their institutions and via their culturally accepted practices, which generally requires seeing these goods as worthy of protection from unjust interference or harms by outsiders. Many paradigmatic

ic cases of just war, such as certain cases of resisting unjust aggression, involve securing or defending a nation's shared impersonal goods, political projects, or institutions. Such wars also generally involve, among other things, protecting conationals' interests and security from threats to their shared political life.<sup>39</sup> All of these elements support the idea that war is, in general, a context in which the conational relationship is salient.

Again, these interests are not always active among conationals who happen to find themselves together. This is what Desert Island demonstrated: examples like this—and others stripped of any relevant contextual features, such as the cases imagined by defenders of the interpersonal cases argument—do not involve a context in which conationals' joint projects, their relationship goods, or their shared history are morally salient. The individuals in Desert Island simply happen to be conationals; their interaction does not occur within a relevant context.

However, when conationals participate in the shared life of their community, or act within or rely upon certain of its institutions in various ways, this changes the context. This is part of what Desert Island Rescue helped to show: interactions mediated through the collective itself can create a context for justified partiality where one would not exist in the absence of that context. As applied to war, the point is that the political and cultural life conationals share is part of what grounds justified partiality; and when these values are threatened, as they typically are in war, or when the security of those participating in our collective life (which aims partly at this very sort of protection) is in jeopardy, the context of the conational relationship becomes salient.<sup>40</sup>

Furthermore, the context of the interaction in certain defensive wars—a paradigmatic case of just war—is importantly different from many other contexts. Many wars of this sort involve an aggressor nation that views the defender nation and its citizens primarily (though of course not exclusively) as targets by

39 This claim is of course clearest to see if one endorses the view that the moral restrictions governing war apply in virtue of one's membership in a particular collective; however, one need not endorse such a view to endorse this claim.

40 Interactions between non-conationals in the course of ordinary political life (e.g., individuals of one nation or state visiting or residing in a foreign nation or state, such as is assumed to be the case in a more concrete version of Lefkowitz's and Fabre's examples) would be mediated by various complicated and contingent facts about what is promised or perhaps expected upon granting entry into a given territory. Thus, even if it is in general true that a particular nation has special reason to prefer the security of its own citizens over nonnationals, nonnationals granted entry (temporary or otherwise) into the boundaries of the nation will usually be doing so under the agreement that they will be, for the proper duration of their stay, treated as though they were conationals. (One could imagine here certain analogies with temporary custody over nonbiological children alongside one's own biological children, for example.)



virtue of their membership in that nation. As we saw in the previous section with Terrorist Vacation, the relevance of the conational relationship changes when the nation and its members are seen by the opposition as a particular target of harm or aggression. This is clearly true of many wars, and one need not look far back in history for examples: the Yugoslav wars of the 1990s are a clear case of hostility on the basis of nationality.<sup>41</sup> When this is true, giving relevant preference to the lives of one's conationals in the course of the war falls within the scope of that relationship.

In sum, the sorts of encounters in war on which we are focused in our present examples indeed fall within the scope of the conational relationship. When we limit ourselves only to cases that fall within the relationship's scope, the conational relationship looks quite similar to the relationship between friends, romantic partners, or parents and children. Some might even think that these reasons are all of roughly similar strength. In a certain sense, then, the analogy that Hurka draws between families and nations is not entirely misguided; however, we must restrict our cases of conational partiality to those that fall within the scope of that relationship for the analogy to work.

This, as I have shown, is the key to rejecting the interpersonal cases argument: we can accept the claim that the conational relationship does not give rise to reasons of partiality in many ordinary interpersonal contexts, while rejecting the claim that is thought to follow from it—namely, that the conational relationship therefore does not give rise to reasons of partiality within the context of war.

One might object here that while this argument does appear to succeed in the case of war, it yields counterintuitive verdicts in other cases, such as the following:

*Hospital Aggressor:* Aggressor threatens to remove Bob's privately owned lifesaving medical device, and the only way you can prevent him from doing so is to throw a grenade that will kill Aggressor and two bystanders. Bob is a conational of yours; the bystanders are not.<sup>42</sup>

Throwing the grenade seems impermissible. But what if the medical device had been provided by a state institution, such as a national health-care program? My view seems to suggest that in the first case, it would be impermissible, because Bob's medical care appears to fall beyond the scope of the conational relationship; however, in the second case, it would be justifiable to throw the grenade, because the device is government issued. And this, one might claim, is implau-

41 For an in-depth discussion of the role of nationality in these wars, see Baker, *The Yugoslav Wars of the 1990s*.

42 Thank you to an anonymous reviewer for raising this objection.

sible. Moreover, this calls into question the basis of the argument as applied to the wartime context.

But as I have argued here, the fact that one's care takes place in the context of a valuable social institution, such as a government-run health-care program, constitutes a significant moral difference; indeed, the same is true in general of cases involving government-issued education (as opposed to private education), government land (versus privately owned property), and so on. Interactions within these contexts are not merely between individuals but conationals as well, and this fact gives rise to special moral permissions, perhaps even duties, to render aid and rescue.

The objector might reply here that who provided a medical device is too flimsy a basis for altering our moral permissions so significantly. Even if we were to grant this point, however, it would not significantly threaten the argument as applied to the wartime case. As I have shown above, the context of war is not the result of a subtle shift in context or a mere technicality. In general, war constitutes a significant shift in the context of the conational relationship: it involves not one but several of the most central institutions that govern the relationship, and it involves not one part of the conational relationship but its very foundations. Even if one is skeptical of a case like Hospital Aggressor, the wartime case is surely different in several morally relevant ways.

#### 4. CONCLUSION

Rejecting the interpersonal-cases argument is an important step toward showing how national partiality can be permitted in war, but it does not fully vindicate Hurka's conclusions on its own. My argument, if successful, only shows that war is a context in which national partiality is salient; it does not show that the conational relationship is powerful enough to justify infringing significant negative duties to others to the extent that Hurka suggests. To show this would require a supplementary argument that focuses not on the applicability of national partiality but on its strength when weighed against competing duties to outsiders. I do not have the space to pursue such an argument here. While many writers on the subject have expressed doubts about national partiality being strong enough to justify serious harming, others believe it can sometimes be justified.<sup>43</sup>

My goal here has not been to defend any specific conclusion about the extent

43 For the former, see McMahan, "Comment on 'Associative Duties and the Ethics of Killing in War'"; Van Goozen, "Harming Civilians and the Associative Duties of Soldiers"; and Betz, "The Priority Problem for the Associativist Theory of Ethics in War." For the latter, see Lazar, "Associative Duties and the Ethics of Killing in War."

to which national partiality can justify actions in war. Rather, my aim was simply to show how the concept of scope-restricted partiality can help us to see why we might agree with the interpersonal arguments, insofar as they show that national partiality is not plausibly justified in many interpersonal cases, and yet reject the claim that it is not applicable in war. The defender of the interpersonal-cases argument might insist that their primary concern is to show that national partiality lacks the strength to justify significant harming and that the interpersonal cases argument was simply their attempt to show that conclusion. But if I am right about scope-restricted national partiality and its application in war, then this approach is unsuccessful in demonstrating that point. In that case, both sides of the debate still need to provide an argument for how we should understand national partiality's strength.<sup>44</sup>

University of Florida  
jeremydavis@ufl.edu

#### REFERENCES

- Baker, Catherine. *The Yugoslav Wars of the 1990s*. London: Palgrave MacMillan, 2015.
- Bazargan-Forward, Saba. "Weighing Lives in War: Foreign vs. Domestic." In *The Cambridge Handbook on the Just War*, edited by Larry May, 186–98. Cambridge: Cambridge University Press, 2018.
- Betz, Adam Thomas. "The Priority Problem for the Associativist Theory of Ethics in War." *Journal of Moral Philosophy* 17, no. 1 (February 2020): 1–32.
- Broad, C. D. *Examination of McTaggart's Philosophy*. Vol. 2. Cambridge: Cambridge University Press, 1938.
- Dworkin, Ronald. *Law's Empire*. Cambridge, MA: Harvard University Press, 1986.
- Fabre, Cécile. *Cosmopolitan War*. Oxford: Oxford University Press, 2012.
- Foot, Philippa. "The Problem of Abortion and the Doctrine of Double Effect." *Oxford Review* 5 (1967): 5–15.

44 Thanks to the audiences in Lyon, Toronto, West Point, Jerusalem, and Kioloa for helpful discussion. Thanks in particular to Kevin Buton-Maquet, Imola Ilyes, and Garrett Cullity for written comments. Thanks as well to Josh Brandt, Steve Coyne, Eric Mathison, Andrew Sepielli, Saba Bazargan-Forward, and Helen Frowe for much helpful discussion. Very special thanks are due to Tom Hurka, Arthur Ripstein, and Julia Nefsky, each of whom helped me considerably to develop the ideas at the core of this essay. This essay also benefited from the support of the Halbert Network Fellowship.

- Gilbert, Margaret. *Living Together: Rationality, Sociality, and Obligation*. Lanham, MD: Rowman and Littlefield, 1996.
- Hardimon, Michael O. "Role Obligations." *Journal of Philosophy* 91, no. 7 (July 1994): 333–63.
- Hurka, Thomas. "The Justification of National Partiality." In *The Morality of Nationalism*, edited by Robert McKim and Jeff McMahan, 139–57. Oxford: Oxford University Press, 1997.
- . "Love and Reasons: The Many Relationships." In *Love, Reason, and Morality*, edited by Esther Kroeker and Katrien Schaubroeck, 163–80. New York: Routledge, 2017.
- . "Proportionality in the Morality of War." *Philosophy and Public Affairs* 33, no. 1 (January 2005): 34–66.
- Kamm, F. M. "Failures of Just War Theory: Terror, Harm, and Justice." *Ethics* 114, no. 4 (July 2004): 650–92.
- Keller, Simon. "Four Theories of Filial Duty." *Philosophical Quarterly* 56, no. 223 (April 2006): 254–74.
- Kolodny, Niko. "Which Relationships Justify Partiality? The Case of Parents and Children." *Philosophy and Public Affairs* 38, no. 1 (Winter 2010): 37–75.
- Lazar, Seth. "Associative Duties and the Ethics of Killing in War." *Journal of Practical Ethics* 1, no. 1 (June 2013): 3–48.
- Lefkowitz, David. "Partiality and Weighing Harm to Noncombatants." *Journal of Moral Philosophy* 6, no. 3 (2009): 298–316.
- Lord, Errol. "Justifying Partiality." *Ethical Theory and Moral Practice* 19, no. 3 (June 2016): 569–90.
- McMahan, Jeff. "Comment on 'Associative Duties and the Ethics of Killing in War.'" *Journal of Practical Ethics* 2, no. 2 (December 2014): 58–68.
- McMahan, Jeff, and Robert McKim. "The Just War and the Gulf War." *Canadian Journal of Philosophy* 23, no. 4 (December 1993): 501–41.
- Miller, David. *On Nationality*. Oxford: Oxford University Press, 1995.
- . "Reasonable Partiality towards Compatriots." *Ethical Theory and Moral Practice* 8, nos. 1–2 (April 2005): 63–81.
- Seglow, Jonathan. *Defending Associative Duties*. New York: Routledge, 2013.
- Shelby, Tommie. *We Who Are Dark: The Philosophical Foundations of Black Solidarity*. Cambridge, MA: Harvard University Press, 2005.
- Stroud, Sarah. "Permissible Partiality, Projects, and Plural Agency." In *Partiality and Impartiality: Morality, Special Relationships, and the Wider World*, edited by Brian Feltham and John Cottingham, 131–49. Oxford: Oxford University Press, 2010.

Swift, Adam, and Harry Brighouse. *Family Values: The Ethics of Parent-Child Relationships*. Princeton: Princeton University Press, 2014.

Thomson, Judith Jarvis. "The Trolley Problem." In *Rights, Restitution and Risk: Essays in Moral Theory*, 94–126. Cambridge, MA: Harvard University Press, 1986.

Van Goozen, Sara. "Harming Civilians and the Associative Duties of Soldiers." *Journal of Applied Philosophy* 35, no. 3 (August 2018): 584–600.

## DEFENDING THE EPISTEMIC CONDITION ON MORAL RESPONSIBILITY

*Martin Montminy*

MORAL RESPONSIBILITY, tradition tells us, requires both a freedom (or control) condition and an epistemic condition. A number of authors have recently challenged this tradition, raising doubts about the epistemic condition. According to these challenges, either moral responsibility does not require such a condition *in addition* to a freedom condition or it does not require an epistemic condition at all. In this paper, I will examine these challenges and argue that they are unsuccessful. In other words, I will argue that moral responsibility does require a nonsuperfluous epistemic condition.

The first two challenges I will consider can be represented as a dilemma. The first horn starts with the thesis that to be blameworthy for an action, the agent must be aware of her wrongdoing. On this view, moral responsibility does involve an epistemic condition. However, some authors argue, this epistemic condition is superfluous. This is because if a person satisfies the condition for acting freely, she must *know* what she is doing. This entails that the epistemic condition is incorporated into the freedom condition: anyone who satisfies the freedom condition automatically satisfies the epistemic condition. The second horn of the dilemma starts with the contention that one may be blameworthy for an unwitting wrongdoing. On this view, no epistemic condition is required for moral responsibility, since a person need not know that she is doing something wrong in order to be blameworthy for her act. Hence, according to this dilemma, either moral responsibility involves no epistemic condition or the epistemic condition on moral responsibility is built into the freedom condition.

I will not take a stance on whether one may be blameworthy for an unwitting wrongdoing. But I will show that both of the challenges against the epistemic condition fail. In section 1, I will argue that although the freedom condition does incorporate some epistemic requirements, these do not include all the epistemic requirements for blameworthiness (and praiseworthiness). Then, in section 2, I will argue that a view that allows for blameworthy unwitting wrongdoings does actually involve an epistemic condition. To put it briefly, the unwitting agent is

blameworthy for an action (or omission) only if she *should* know, in a sense to be explained, that that action (omission) is wrong. In section 3, I will identify more specifically what epistemic requirements the freedom condition on moral responsibility involves. This will provide further support for the traditional view that moral responsibility requires a nonsuperfluous epistemic condition.

Section 4 will be concerned with the third challenge to the epistemic condition. According to standard formulations of the quality of will view, a person is blameworthy for a wrong action just in case the action manifests ill will or a lack of goodwill. Blameworthiness does not involve any epistemic condition. I will argue that to spell out what it means for an action to *manifest* a problematic quality of will, one must invoke some epistemic requirements.

#### 1. DOES THE FREEDOM CONDITION INCORPORATE THE EPISTEMIC CONDITION?

John Fischer and Mark Ravizza advocate the traditional view according to which moral responsibility involves not only a freedom condition but also an epistemic condition. To support their position, they invoke an instance of *nonculpable ignorance*:

*Dead Kitten*: Kit is backing his car out of his garage. Kit is not under the influence, and his judgment is not addled in any way. Through no fault of his own, he is unaware that a kitten is lying right outside the garage door. Unfortunately, Kit's car is not equipped with a backup camera. Kit drives over the kitten while slowly backing his car out of his garage and kills it.<sup>1</sup>

Intuitively, Fischer and Ravizza point out, Kit is not morally responsible for killing the kitten. This is because he is unaware of the kitten's presence outside the garage, and by assumption, his ignorance is blameless. However, it seems that Kit acts freely when he is backing up the car: there is nothing wrong with his state of mind and he is in control of the car. Ignorance thus seems like an excusing condition of a different kind than lack of freedom. Cases of nonculpable ignorance such as *Dead Kitten* suggest that moral responsibility involves an epistemic condition in addition to a freedom condition.

Alfred Mele challenges this suggestion.<sup>2</sup> First, he points out, the claim that Kit *freely* kills the kitten is highly counterintuitive. Plausibly, Kit freely holds the wheel in a certain way, freely presses the gas pedal, and freely moves the car back-

1 This story is adapted from both Fischer and Ravizza, *Responsibility and Control*, 12; and Mele, "Moral Responsibility for Actions," 104.

2 Mele, "Moral Responsibility for Actions," 104.

ward. And he is morally responsible for these actions.<sup>3</sup> By contrast, he does not freely run over the kitten. Hence, one could argue, the reason Kit is not morally responsible for running over the kitten is that he does not perform this action freely. We can thus explain Kit's blamelessness without invoking an epistemic condition: the freedom condition suffices.

This brief explanation rests on the plausible principle that a person cannot perform an action freely unless she knows what she is doing. In other words, the freedom condition on action incorporates an epistemic requirement: a person who satisfies the freedom condition with respect to action *A* automatically satisfies an epistemic condition with respect to *A*. The epistemic condition on moral responsibility thus seems superfluous.

Here is another way to reach the same conclusion. Mele considers the principle, which he attributes to Fischer and Ravizza, that nothing is a free action unless it is an intentional action.<sup>4</sup> In other words, free entails intentional, or FI:

FI: Every free action is an intentional action.

If FI is correct, then any epistemic requirement on intentional action is also a requirement on free action. Now, as Elizabeth Anscombe contends, when we do something intentionally, we know that we are doing it.<sup>5</sup> Anscombe's contention, combined with FI, entails that a free action is a clear-eyed action:

FK: One performs *A* freely only if one knows that one performs *A*.

We may invoke FK to conclude that the reason Kit is not blameworthy for running over the kitten is that he did not perform this action freely. The absence of knowledge entails the absence of freedom. Kit is blameless because the freedom condition is not satisfied. Therefore, ignorance excuses because lack of freedom excuses.

Mele asks, "What epistemic requirements for being morally responsible for performing an action *A* are not also requirements for freely performing *A*?" His answer is "I do not know."<sup>6</sup> Like Mele, Neil Levy considers Fischer and Ravizza's

3 As Mele points out, it is a contentious matter whether moving the car backward in a certain way and killing the kitten are different actions or, according to a Davidsonian coarse-grained view, the same action under different descriptions. Like Mele, I wish to remain neutral on the individuation of actions. However, in what follows, I will use the simpler language of the fine-grained view. Translation to a Davidsonian language should be straightforward.

4 Mele, "Moral Responsibility for Actions," 106; Fischer and Ravizza, *Responsibility and Control*, 64.

5 Anscombe, *Intention*, 13–14.

6 Mele, "Moral Responsibility for Actions," 108.



claim that moral responsibility involves an epistemic condition that is distinct from the freedom condition:

Fischer and Ravizza, for instance focus on the control condition, *rather than* the epistemic condition. This is a mistake, I shall argue: the epistemic condition isn't independent of the control condition but built right into it. . . . Possession of the kind of control that matters in the debate, the kind of control that (allegedly) justifies attributions of moral responsibility, requires satisfaction of a demanding set of epistemic conditions.<sup>7</sup>

Levy imagines the following case:

*Peanut Butter*: Betty has a peanut allergy, but Grandfather does not know that. Moreover, Grandfather's ignorance is nonculpable: he was never informed that Betty is allergic to peanuts and has no reason to suspect that she suffers from this condition. At lunch, Grandfather feeds Betty peanut butter. This produces an anaphylactic reaction, and Betty must be brought to the emergency room.<sup>8</sup>

Because of his ignorance, Grandfather is not blameworthy for Betty's anaphylactic reaction. However, Levy points out, he certainly does not intentionally or freely cause that reaction. He thus lacks control over that: intuitively, he would have had control over the reaction only if he had known about Betty's allergy. Hence, once again, to account for Grandfather's blamelessness, there is no need to appeal to an epistemic condition. We can simply invoke the control (or freedom) condition on moral responsibility, which incorporates the epistemic condition.

Mele's and Levy's objections rest on a mistake, though. Grandfather freely and intentionally gives peanut butter to Betty. Given  $\text{FK}$ , this means that he knows that he is giving her peanut butter. But giving peanut butter to Betty is morally wrong, since it will (likely) produce an anaphylactic reaction in her. Why is Grandfather blameless for serving peanut butter to Betty, then? Because he does not know that giving Betty peanut butter is morally wrong. (Alternatively, the answer could be that he does not know that giving Betty peanut butter will harm her. I will come back to this point in section 4.) In other words,

7 Levy, *Hard Luck*, 110–11. Levy's criticism of Fischer and Ravizza is ambiguous. He may be interpreted as holding that "the" epistemic condition on moral responsibility is, as he puts it, "built right into" the freedom condition. This is how I read him here. He could also be making the weaker claim that satisfying the freedom condition "requires satisfaction of a demanding set of epistemic conditions." This weaker claim is true; however, it is an unfair objection to Fischer and Ravizza, since, as we will see in section 3, they impose several epistemic constraints on free action.

8 Levy, *Hard Luck*, 113.

Grandfather is not blameworthy for serving Betty peanut butter because he does not satisfy the epistemic condition with respect to this action, even though he performs this action freely.

The point is that the epistemic condition on moral responsibility for an action concerns not only what the action is but also its moral significance. In other words, to be blameworthy for a wrongdoing *A*, a person must know not just that he is performing *A* but also that *A* is morally wrong. And, as we saw, satisfying the freedom (control) condition with respect to action *A* only requires that the agent know that he is performing *A*. For this reason, the freedom (control) condition does not incorporate the epistemic condition.

Consider Dead Kitten again. Kit does not freely and intentionally kill the kitten. But he does back his car out of his garage at a certain time and in a certain manner. Call this action *backing out*, for short. Given the presence of the kitten, backing out is the wrong thing to do. Now, Kit backs out freely and intentionally. He is not blameworthy for backing out, since he is not aware that backing out is morally wrong. An epistemic condition that is not entailed by the freedom condition is thus needed to account for the fact that Kit is blameless for backing out.

Mele's and Levy's arguments would go through if they were restricted to actions such as wrongfully backing out or wrongfully serving peanut butter. Suppose Grandfather freely (and intentionally) wrongfully serves peanut butter to Betty. Grandfather has control over his wrongfully serving peanut butter. This, in turn, very plausibly entails that he knows that he is wrongfully serving peanut butter. Here, the freedom condition on wrongfully serving peanut butter does incorporate the epistemic condition on that action. Unfortunately, as we saw, this argument does not extend to backing out or serving peanut butter.

It is worth considering a similar argument that focuses on praiseworthiness. Paulina Sliwa contends that "moral responsibility inherits its epistemic condition from the epistemic condition on intentional action."<sup>9</sup> Sliwa first points out that the success of an intentional action *A* should not be accidental. One may intend to do *A* and succeed, but if one's success is a matter of chance, then, intuitively, one does not do *A* intentionally. The outcome of an intentional action, Sliwa remarks, should result from the exercise of agential control. She characterizes this control as *know-how*. According to Sliwa, one intentionally does *A* only if *A* results from one's intention to perform *A* and one's knowledge of how to do *A*. Applied to right actions, this principle entails that "intentionally doing the right thing requires both an intention to do what's right and knowledge of how to do the right thing."<sup>10</sup> Sliwa then turns to praiseworthiness. She contends that

9 Sliwa, "On Knowing What's Right and Being Responsible for It," 127.

10 Sliwa, "On Knowing What's Right and Being Responsible for It," 131.

praiseworthiness for a right action requires intentionally doing the right thing. Therefore, moral knowledge is a necessary condition on praiseworthiness for a right action because it is a necessary condition on intentionally acting rightly.

Let us grant Sliwa's conclusion that one is praiseworthy for an action only if one knows that one is acting rightly. Does this conclusion entail that praiseworthiness inherits its epistemic condition from the epistemic condition on intentional action? No. Suppose Zara donates a certain amount of money to a charity. Call Zara's action *donating*, for short. According to Sliwa's condition on praiseworthiness, Zara is praiseworthy for her action only if she intentionally does the right thing. And given Sliwa's condition on intentional action, this entails that Zara is praiseworthy for her action only if she knows that she is acting rightly. This means that Zara is praiseworthy for donating only if she knows that donating is the right thing to do or, as I will write, only if she knows that she is *donating rightly*. But intentionally donating is not the same as intentionally donating rightly. Only the latter requires knowledge that donating is the right thing to do. Hence, while the epistemic condition on praiseworthiness for donating requires that Zara know that she is donating rightly, the epistemic condition on intentionally donating only requires that Zara know that she is donating. This means that praiseworthiness for donating does not inherit its epistemic condition from the epistemic condition on intentionally donating. Sliwa's thesis is true only of actions such as donating rightly: praiseworthiness for donating rightly does inherit its epistemic condition from the epistemic condition on intentionally donating rightly.

Simply put, my argument is as follows. Not every right action is transparently right, and not every wrong action is transparently wrong. Plausibly, as we have seen, a person who freely and intentionally does *A* knows that she is doing *A*. But this does not entail that she knows whether *A* is right or wrong. This means that the epistemic requirements for being morally responsible for doing *A* are not also requirements for freely and intentionally doing *A*.

So far, I have assumed that the epistemic condition on blameworthiness (praiseworthiness) for an action requires knowledge that the action is wrong (right). This assumption is contentious. In the next section, I will consider a way in which the condition may be weakened with respect to blameworthiness. In section 4, I will mention an alternative way to construe the content of the knowledge attitude. But I should also point out that it is not obvious that the attitude involved in the epistemic condition should be knowledge. Some have argued for other attitudes, such as true belief, justified belief, or mere belief.<sup>11</sup> In the rest of the essay, I will talk as if knowledge is the attitude invoked by the epistemic condition, but substituting another attitude for knowledge would not affect my arguments.

11 See Rudy-Hiller, "The Epistemic Condition for Moral Responsibility," for a useful overview.

## 2. UNWITTING WRONGDOINGS

According to the epistemic condition discussed in the previous section, a person is blameworthy for a wrong action *A* only if she knows that *A* is wrong. Consider first a flawed counterexample to this condition.

*Asleep at the Wheel*: Shasta is very eager to reach her destination and has been driving her car for over ten hours. Her eyelids are droopy and her head starts to nod. She ignores these symptoms and keeps driving. Soon, she falls asleep and collides with another car.

Clearly, when she is asleep, Shasta has no awareness of the other car. However, intuitively, she is to blame for the collision.

*Asleep at the Wheel* does not threaten the epistemic condition, though. This is because Shasta is *derivatively* (or *indirectly*) blameworthy for colliding with the other vehicle. Her blameworthiness for the collision derives from her direct blameworthiness for driving while feeling sleepy. The epistemic condition, it should be clear, is meant to be a condition on *direct* moral responsibility and blameworthiness. Plausibly, Shasta is directly blameworthy for driving while feeling sleepy. Moreover, she plausibly satisfies the epistemic condition regarding that act, since she is plausibly aware that it is wrong. Hence, *Asleep at the Wheel* does not threaten the epistemic condition.

Here is a case, imagined by George Sher, that more clearly challenges the epistemic condition.<sup>12</sup>

*Hot Dog*: Alessandra is picking up her children at their elementary school. Although it is hot, she leaves Sheba, the family dog, in the car while she goes to gather her children. Since the pickup is always quick, this has never been a problem. This time, however, Alessandra has several conversations with other parents and completely forgets about Sheba. When she and her children finally come back to the parking lot, they find Sheba unconscious from heat prostration.

According to Sher, Alessandra is morally responsible (and blameworthy) for leaving Sheba in the hot car for so long. More specifically, Alessandra is *directly* blameworthy for that omission. (From now on, when the context makes it clear, I will omit “directly” before “blameworthy.”) *Hot Dog* is an instance of a *blame-*

12 Sher, *Who Knew?* 24. For similar cases, see Amaya, “Slips”; Clarke, *Omissions*, 164; Harman, “Does Moral Ignorance Exculpate?” 463; and Angela Smith, “Responsibility for Attitudes,” 236.

worthy unwitting omission: Alessandra does not knowingly leave Sheba in the hot car. Hence, she is blameworthy for doing something she does not know is wrong.

Let us grant Sher's assessment, and see what it entails regarding the epistemic condition. As Sher points out, Hot Dog is a counterexample to the *searchlight view*, a view according to which a person is responsible only for the features of her acts that she is aware of.<sup>13</sup> However, Hot Dog does not show that responsibility involves no epistemic requirement. Discussions of the epistemic condition often construe this condition as *disjunctive*: a person is blameworthy for a wrongdoing only if she knows or *should know* (or *is reasonably expected to know*) that she is doing the wrong thing.<sup>14</sup>

There are several ways to specify what "should know" means. For example, it can mean that the person was in a position, *prior to her action*, to investigate the matter. A paramedic who is unable to use a defibrillator when attending a patient in cardiac arrest is blameworthy for her inaction. This is because she should know how to operate the device, given her training. If her ignorance is due to her inattention in class or her failure to study, then she is *derivatively* blameworthy for her incompetent treatment of the patient. Since our interest is in direct blameworthiness, I will set aside this particular understanding of "should know."

Hot Dog provides us with another way to understand "should know." Alessandra should know that Sheba is at risk of heat prostration, given her evidence (the high temperature, the fact that Sheba is in the car, and so on) and her cognitive capacities. Alessandra is at fault because, while she is talking with other parents, she is capable of remembering Sheba's situation. Her forgetting, as Sher puts it, is *substandard*: given her cognitive capacities, Alessandra should have remembered.<sup>15</sup> In other words, Alessandra's forgetting is a type of *underperformance*: she does not perform according to her cognitive abilities.<sup>16</sup>

More recently, Sher has proposed a statement of the epistemic condition that accommodates his case: "When a morally ignorant wrongdoer satisfies all the non-epistemic conditions for blameworthiness, he is blameworthy for acting wrongly if, but only if, he was *at least in a position to recognize* that his act was wrong when he performed it."<sup>17</sup> Clearly, Sher's epistemic condition is not automatically satisfied when the freedom condition is. We may assume that Alessan-

13 Sher, *Who Knew?* 6.

14 See, among others, Fischer and Ravizza, *Responsibility and Control*, 12; Ginet, "The Epistemic Requirements for Moral Responsibility"; and Vargas, *Building Better Beings*, 201–2.

15 Sher, *Who Knew?* 110.

16 Montminy, "Doing One's Reasonable Best"; Raz, "Being in the World"; Michael Smith, "Rational Capacities."

17 Sher, "Blame and Moral Ignorance," 101.

dra freely stays in the school to talk to other parents for several minutes. She is blameworthy for this action, because she is in a position to know that it is wrong.

A possible response is worth considering. One may point out that it would not be unreasonable to broaden the freedom condition similarly to the way in which we just broadened the epistemic condition. In a sense, Alessandra was free to come back to the car more quickly. She had that freedom: she just failed to exercise it appropriately. As Randolph Clarke remarks,

Even if, in a case of unwitting omission, the agent doesn't *exercise* control with respect to whether she performs the omitted action, she might *have* control over whether she performs that action, for she might *be able to exercise*, and *have the opportunity to exercise*, such control—she might be able, and have the opportunity, to perform the action in question. She might thus be *free* to perform it, even if there's nothing that is her omission and is an exercise of freedom.<sup>18</sup>

According to the broader freedom condition, an action is free either if the agent performed it freely or if she was free to perform it. Arguably, Alessandra was free to come back to her car in time. Moreover, although Alessandra did not intentionally leave the dog in the hot car, she was able to intentionally come back to the car. And if she had intentionally come back to her car, she would have known that she was going back to her car.

The broader freedom condition thus incorporates an epistemic condition. Does it incorporate the epistemic condition on blameworthiness? No. If Alessandra had freely come back to her car, she would have known that she was going back to her car. But Alessandra's capacity to form this knowledge does not give her knowledge of the moral significance of either her current or her alternative action. Generally speaking, according to the broader freedom condition, a person is blameworthy for not doing *A* only if she was free to do *A*. Based on plausible assumptions, this means that a person is blameworthy for not doing *A* only if she had the capacity to know, while doing *A*, that she was doing *A*. One may possess this capacity without satisfying the epistemic condition on blameworthiness. According to the latter, the person should be able to know, through the exercise of her cognitive abilities, that not doing *A* is wrong. Hence, one may satisfy the broader freedom condition with respect to an omission without satisfying the epistemic condition on blameworthiness for that omission.

Following Clarke, let us call the view presented in this section the *ability view*

18 Clarke, "Blameworthiness and Unwitting Omissions," 70. See also Clarke, *Omissions*, ch. 5, and "Ignorance, Revision, and Commonsense," 249.

(hereafter AV).<sup>19</sup> A morally responsible person has cognitive and volitional abilities, and according to AV, she is blameworthy for a wrongdoing *A* only if her doing *A* results from a substandard failure to exercise these abilities. (Recall that a failure is substandard when the agent fails to exercise her own abilities successfully. In other words, the agent fails to perform according to her abilities.) According to AV, blameworthiness has two sources: it may be due to a substandard failure to exercise one's motivational or volitional abilities to do the right thing (clear-eyed wrongdoing); it may also be due to a substandard failure to recognize what the right thing to do is (unwitting wrongdoing).

### 3. THE EPISTEMIC REQUIREMENTS ON FREEDOM

As we saw in section 1, both Mele's and Levy's arguments target Fischer and Ravizza's account of moral responsibility. Mele writes, "According to Fischer and Ravizza, what epistemic requirements for being morally responsible for performing an action *A* are not also requirements for freely performing *A*? My answer at this point is straightforward, even if it is unsatisfying: I do not know."<sup>20</sup> And Levy contends that Fischer and Ravizza mistakenly attempt to treat the freedom condition independently of the epistemic condition. Now, both Mele and Levy correctly point out that freedom does incorporate some epistemic requirements. In this section, I will try to specify this point further, focusing on Fischer and Ravizza's account of the freedom condition. This will provide further support for the claim that a nonsuperfluous epistemic condition on moral responsibility is needed.

Fischer and Ravizza characterize the freedom condition on moral responsibility as *guidance control*. They identify guidance control over an action *A* with *A*'s resulting from the agent's own reasons-responsiveness.<sup>21</sup> Reasons-responsiveness consists in two mechanisms: a *reasons-receptive* mechanism, which is in charge of recognizing the reasons for an action, and a *reasons-reactive* mechanism, whose role is to act on those reasons in a given circumstance. Reasons-receptivity is thus a kind of cognitive ability, and reasons-reactivity, a kind of volitional ability.<sup>22</sup>

19 Clarke, "Blameworthiness and Unwitting Omissions."

20 Mele, "Moral Responsibility for Actions," 108.

21 Fischer and Ravizza, *Responsibility and Control*, chs. 2–3. To exclude cases of manipulation, Fischer and Ravizza insist that the mechanisms should be the agent's *own*, and then they analyze mechanism ownership in terms of *taking responsibility*. I will not explore this part of their account.

22 I will often talk of people's rather than mechanisms' reasons-responsiveness. Fischer and Ravizza's focus on mechanisms is motivated by considerations about Frankfurt-style cases. This will not be a concern here, as I will assume that Frankfurt devices are always absent.

According to Fischer and Ravizza, a person performs an action *A* freely on a given occasion only if *A* results from reasons-responsive mechanisms. These mechanisms should be such that in a reasonable proportion of nearby possible scenarios in which there are sufficient reasons to do otherwise, the person's reasons-receptive mechanism recognizes these reasons and the person's reasons-reactive mechanism chooses otherwise based on these reasons. (Given our purposes, it is not crucial to specify the threshold marked by "reasonable proportion.") But this is not all. To be free, *A* must be produced in the right way by the mechanisms. This means that the agent must act based on a reason *R*. Four points are worth emphasizing. First, *R* need not be a sufficient, or even a good, reason for *A*. To perform *A* freely, the agent must have the capacity to recognize sufficient reasons, but she may be mistaken that *R* is sufficient for *A* on this particular occasion. For example, a parent may freely spank a child to discipline him based on a mistaken belief that it is the right thing to do. Second, *R* need not be a *moral* reason: acting freely does not require acting based on moral reasons. A free action may be based on a person's (nonmoral) values. For example, one can freely eat a piece of cake because one enjoys the taste. Third, one may do *A* based on *R*, even though one takes *R*\* to be a sufficient reason for an alternative course of action *B*. To freely do *A*, an agent must have the capacity to do otherwise based on sufficient reasons to do otherwise, but the agent need not have a perfect record. Clear-eyed wrongdoings illustrate this point: even though the agent knows that there is a sufficient moral reason to do *B* rather than *A* because *B* is morally right and *A* is not, she may still do *A* freely. Fourth, according to principle FK of section 1, to do *A* freely, the agent must know that she is doing *A*.

The freedom condition with respect to action *A* thus incorporates some epistemic requirements, but it does not include the requirement that the agent know or should know about the moral significance of *A*. Hence, moral responsibility for *A* involves an epistemic condition that goes beyond what the freedom condition requires. I will illustrate this point with some examples, starting with a praiseworthy action. Suppose Meili rescues a trapped child from a burning house. Call Meili's action *rescuing*, for short. To rescue freely, Meili must satisfy different epistemic requirements: (1) she must have the ability to recognize sufficient reasons to do otherwise in a reasonable proportion of nearby possible scenarios, (2) she must consider (explicitly or implicitly) that there is a reason for rescuing, and (3) she must know that she is rescuing. Note that 2 does not require Meili's reason for rescuing to be moral. She may rescue for a monetary reward, for example. This would not make her action unfree. Clearly, 1–3 do not require Meili to know that *rescuing* is the right thing to do. But as we saw in section 1, to be praiseworthy for her action, Meili must know that rescuing is right or



recognize that there is a morally sufficient reason for rescuing. Therefore, praiseworthiness for rescuing involves an epistemic condition that is additional to the freedom condition.

The conditions on blameworthiness for a wrongdoing are a little more complicated, because we may allow for blameworthy unwitting wrongdoings in addition to blameworthy clear-eyed wrongdoings. Let us start with an example of the latter. Suppose a hired assassin kills an innocent person for money. Call this action *killing*, for short. In this case, the assassin freely kills while knowing that the right course of action would be not killing. In other words, he recognizes that there is sufficient moral reason for not killing, but he nevertheless kills. The assassin kills not based on the recognition of sufficient moral reasons, but based on the recognition of some selfish reason—that is, his desire for money. Clearly, this fact does not entail that his action is not free. To kill freely, the assassin must satisfy the following epistemic requirements: (1) he must have the ability to recognize sufficient reasons to do otherwise in a reasonable proportion of nearby possible scenarios, (2) he must consider that there is a reason for killing, and (3) he must know that he is killing. But the epistemic condition on blameworthiness for killing requires that the assassin recognize some sufficient *moral* reason for not killing or that he recognize that killing is wrong. Once again, the freedom condition on blameworthiness does not incorporate the epistemic condition.

Let us turn to Hot Dog, a case of unwitting wrongdoing. Let us call *chatting* Alessandra's free action of chatting for a significant period of time with other parents while in the school building. In this case, chatting is the wrong thing to do: Alessandra should go back to the parking lot and release Sheba from the hot car (hereafter *going back*). Alessandra does not realize that chatting is the wrong thing to do and that there is a sufficient moral reason for going back. However, we may suppose that Alessandra chats based on the recognition of some reason for chatting, say, to cultivate her friendly rapport with other parents. To chat freely, Alessandra must satisfy the following epistemic requirements: (1) Alessandra must have the ability to recognize sufficient reasons to do otherwise in a reasonable proportion of nearby possible scenarios, (2) she must consider that there is a reason for chatting, and (3) she must know that she is chatting. Now, recall that according to AV, to be blameworthy for chatting (and for not going back), Alessandra must have the ability to recognize that there is a sufficient moral reason for not chatting (and a sufficient moral reason for going back). Clearly, the fact that Alessandra satisfies 1–3 does not entail that she satisfies this epistemic condition on blameworthiness. Hence, the epistemic requirements entailed by the freedom condition do not incorporate the epistemic condition on blameworthiness in this case either.

Now, I need to consider one complication. Toward the end of their discussion, Fischer and Ravizza strengthen their account of guidance control by requiring that the agent be receptive not just to reasons in general but also to *moral* reasons.<sup>23</sup> This, they point out, allows them to hold that intelligent animals, young children, or psychopaths who are unable to appreciate moral reasons are not morally responsible for their actions.<sup>24</sup> I should first note that although receptivity to moral reasons is very plausibly a condition on *moral responsibility*, it does not strike me as a plausible condition on *freedom*. A person who is unable to grasp moral reasons should not be held morally responsible for his actions; however, such a person could still act freely it seems. I would instead hold that the capacity to grasp moral reasons is an epistemic condition on moral responsibility that is additional to the epistemic requirements entailed by freedom. At any rate, let us assume that Fischer and Ravizza are right and that the freedom condition on moral responsibility does require the capacity to recognize moral reasons. One may possess this capacity without satisfying the epistemic condition on blameworthiness. Consider Alessandra again. Fischer and Ravizza's additional epistemic requirement entails that she cannot be blameworthy for anything unless she has the ability to recognize moral reasons. We can safely assume that Alessandra, being a normal human adult, does meet this epistemic requirement. But this does not entail that she possesses the ability to recognize that chatting is the wrong thing to do in her circumstances and that she should go back instead. Hence, an epistemic condition is still required in addition to Fischer and Ravizza's strengthened freedom condition.

#### 4. QUALITY OF WILL

Let us now consider the third challenge against the epistemic condition on mor-

23 Fischer and Ravizza, *Responsibility and Control*, 76–81.

24 Surprisingly, Fischer and Ravizza claim that moral responsibility does not require that the agent be *reactive* to moral reasons (*Responsibility and Control*, 79). According to them, a person may be morally responsible for his actions even though he is completely unable to translate moral reasons into action. Two remarks. First, Fischer and Ravizza's position is implausible: it seems that a person who is unable to react to moral reasons has an excuse for his wrongdoings. Second, Fischer and Ravizza's position appears to contradict a principle they invoke in defense of their contention that weak reactivity to reasons is all the reactivity required for moral responsibility. According to what they call a *fundamental intuition*, reactivity is "all of a piece." By that, they mean that if an agent can react to *some* incentive to do other than he actually does, then he can react to *any* incentive to do otherwise (73). Hence, according to Fischer and Ravizza's fundamental intuition, any person with weak reactivity should be able to translate moral reasons into action. (Thanks to an anonymous reviewer for pointing this out.)

al responsibility. According to the quality of will view (hereafter QW), a person is blameworthy for a wrong action just in case the action manifests ill will or a lack of goodwill. A person who hurts others out of hatred or sadism manifests ill will. Culpability may also be due to a lack of goodwill, or moral indifference. Someone who steals or lies for purely selfish reasons manifests moral indifference.

QW is a family of views. Some of its proponents reject the possibility of blameworthy unwitting wrongdoings. Michael McKenna, an advocate of QW, suggests that a person is blameworthy for an action only if she knows that her action is wrong.<sup>25</sup> On this version of QW, only clear-eyed wrongdoings may be blameworthy. This brand of QW thus explicitly admits an epistemic condition. Given that my interest is in challenges to this condition, I will disregard this version of QW and focus solely on versions of QW that allow for blameworthy unwitting wrongdoing.<sup>26</sup>

When a person is blameworthy for a clear-eyed wrongdoing, she manifests a bad quality of will roughly as follows. The person has problematic desires (or pro-attitudes)—for example, a selfish desire or a desire for the wrong. Since the person's problematic desires are stronger than her desire for the right, they bring about (in a nondeviant way) the wrongdoing. The person's action thus manifests a bad quality of will. Suppose that Fernanda detests her colleague Violeta. Because of that, her desire to hurt Violeta is stronger than her concern for Violeta's welfare. During lunch break, Fernanda's problematic desire causes her to cruelly tease Violeta. Fernanda is blameworthy for her action because it manifests ill will toward Violeta.

Things are very different in cases of blameworthy unwitting wrongdoings. In such cases, the wrongdoer either harbors a certain dose of ill will or lacks the proper amount of goodwill. This problematic quality of will somehow affects his ability to acquire the relevant knowledge at the right time. This, in turn, leads him to unwittingly do the wrong thing. Nomy Arpaly and Timothy Schroeder, two proponents of QW, discuss the case of Victor, a professor who, unbeknownst to him, is prejudiced against students who wear their baseball caps backward.<sup>27</sup> While grading essays, Victor unknowingly gives a student an unfair grade because he remembers the student as wearing his baseball caps backward. Victor is

25 McKenna, *Conversation and Responsibility*, 15, 61.

26 See, among others, Arpaly, *Unprincipled Virtue*, ch. 3; Arpaly and Schroeder, *In Praise of Desire*; Björnsson, "Explaining (Away) the Epistemic Condition on Moral Responsibility"; Talbert, "Akrasia, Awareness, and Blameworthiness"; and Talbert, "Omission and Attribution Error."

27 Arpaly and Schroeder, *In Praise of Desire*, 238.

blameworthy for giving the student an unfair grade because his action manifests ill will toward the student.

It is worth noting that for this kind of blameworthy unwitting wrongdoing to occur, the agent's problematic desires need not be stronger than his desire for the good. It may well be that if he came to realize that the grade he gave to the student is unfair (perhaps by comparing his essay to another of similar quality to which he gave a better grade), Victor would immediately adjust his grade, because his goodwill toward the student is stronger than his ill will toward him.

This shows that the process by which a person may manifest a bad quality of will in an unwitting wrongdoing is very different from the process by which he may manifest ill will in a clear-eyed wrongdoing. What the two instances of "manifesting" a bad quality of will have in common is that the person's problematic will is a cause of the person's wrongdoing. Plausibly, in both cases, the causal chain from problematic will to wrongdoing ought to be *nondeviant*. But what makes nondeviant a causal process from an objectionable desire to a clear-eyed wrongdoing has probably very little in common with what makes nondeviant a causal process from an objectionable desire to a failure to form the appropriate cognitive state at the right time.

To specify further what manifesting a problematic will amounts to in the two types of cases, we must invoke an epistemic condition.<sup>28</sup> I will now argue that *qw*, like *av*, is committed to a disjunctive epistemic condition on blameworthiness: a person is blameworthy for a wrongdoing only if she knows or should know that she is doing the wrong thing. As we will see, the key difference between *qw* and *av* concerns how to construe "should know."

Consider *qw*'s treatment of Hot Dog. According to *qw*, Alessandra is to blame for leaving Sheba in the hot car only if her forgetting is due to ill will or a lack of goodwill. Otherwise, her omission would not manifest a bad quality of will. *qw* thus rejects *av*'s epistemic condition, according to which a person is to blame for an unwitting wrongdoing if, given her abilities, she was in a position to recognize that her act was wrong when she performed it. Against this view, Matthew Talbert, a proponent of *qw*, writes: "Blame might be an apt response if we filled in the details so that it was clear that Alessandra's forgetting stemmed from faulty concern for Sheba's welfare, but if the explanation for Alessandra's behavior doesn't make reference to something like this, then I don't see how the morally offended responses involved in blame will have much purchase."<sup>29</sup>

28 Very plausibly, we must also invoke a freedom condition not too different from the one presented in the previous section. See, for instance, Arpaly, *Unprincipled Virtue*, ch. 4.

29 Talbert, "Akrasia, Awareness, and Blameworthiness," 57. For a similar point, see Björnsson, "Explaining (Away) the Epistemic Condition on Moral Responsibility."

Suppose that Alessandra forgot about Sheba because of her lack of concern for her. Her forgetting is thus causally explained by her objectionable quality of will. This means that if she had had an appropriate degree of concern for her dog, Alessandra would have remembered her. Therefore, Alessandra is to blame for her omission, because she should have known that she was doing something wrong, in the sense that if she had had an adequate quality of will, she would have known that.<sup>30</sup>

This epistemic condition can be further supported by considering a slightly different version of the story, which I will call *Hot Dog\**. In this version, while in the school building, Alessandra encounters a series of extremely stressful events such as an altercation between parents, a sick schoolchild treated by a nurse, and a conversation about a bomb threat that occurred earlier in the day. In *Hot Dog\**, Alessandra is not to blame for leaving Sheba in the car, even if her concern for Sheba's welfare is faulty. This is because regardless of the quality of her will, she would have forgotten about Sheba, given her exceptionally stressful circumstances.<sup>31</sup> According to QW, a blameworthy person's problematic will must be a *difference maker*: Alessandra's lack of concern for Sheba must causally explain her forgetting. QW thus involves the following epistemic condition: a person is blameworthy for an unwitting wrongdoing only if she would know that she is doing the wrong thing if she had an appropriate quality of will.

At this point, it is worth mentioning an area of contention among proponents of QW. Michael Smith distinguishes between a *de re* and a *de dicto* desire for the right.<sup>32</sup> Desiring the right *de re* (that is, desiring to do what is actually right) is not the same as desiring the right *de dicto* (that is, desiring to do what is right whatever that turns out to be). A person with goodwill, Smith insists, should care about equality, justice, and the well-being of others. A concern for the right as such does not make a person good; it is, as Smith puts it, a form of fetishism. Nomy Arpaly, Timothy Schroeder, and Julia Markovits concur with this position.<sup>33</sup> Zoë Johnson King and Paulina Sliwa favor a *de dicto* condition

30 Here, I am employing a simple counterfactual analysis of causation à la Lewis, "Causation": Alessandra's bad quality of will caused her forgetting just in case had she had an adequate quality of will, she would have remembered. This simple analysis would clearly need to be refined to accommodate cases of overdetermination, preemption, and the like; however, it will suffice for our purposes.

31 Both Björnsson and Talbert acknowledge this point. See Björnsson, "Explaining (Away) the Epistemic Condition on Moral Responsibility," 153; Talbert, "Omission and Attribution Error," 23–24.

32 Michael Smith, *The Moral Problem*, 73–76.

33 Arpaly and Schroeder, *In Praise of Desire*, 164–67; Markovitz, "Acting for the Right Reasons," 218–19.

on goodwill.<sup>34</sup> This debate is relevant to the *content* of the attitude invoked in the epistemic condition. Blameworthiness for an unwitting wrongdoing *A* may require that a person with an adequate quality of will would know that *A* is wrong (*de dicto*). This is how I have characterized the epistemic condition so far. But this condition could instead require that a person with an adequate quality of will would know that *A* has such-and-such wrong-making features (*de re*). In Hot Dog, the condition would hold that if she had goodwill, Alessandra would know that her leaving Sheba in the hot car is harming her. Although I will continue to interpret the epistemic condition as *de dicto*, my discussion is meant to be neutral between the two readings.

QW's treatment of Peanut Butter would also invoke the epistemic condition I just described. Grandfather does not know that feeding peanut butter to Betty is wrong. He is blameworthy for serving peanut butter to Betty only if he *should know* that this is wrong. According to QW, Grandfather should know about the wrongness of feeding peanut butter to Betty just in case he would know that feeding peanut butter to Betty is wrong if he had an appropriate quality of will. If Grandfather's ignorance is due to an inadequate degree of concern for Betty, then he is blameworthy for his action. In such a case, he should know that his action is wrong. But in presenting the case, we have assumed that Grandfather was not at fault for his ignorance: this means that regardless of his quality of will, it would have been impossible for him to know about Betty's allergy.

Now, it would be useful to provide an account of blameworthiness that is neutral between AV and QW. Both views admit two types of blameworthy actions: clear-eyed wrongdoings and unwitting wrongdoings. A person is blameworthy for a clear-eyed wrongdoing only if her failure to refrain from doing the wrong thing is substandard. A person is blameworthy for an unwitting wrongdoing only if her ignorance, or failure to recognize the wrongness of her action, is substandard. AV and QW agree on both of these conditions but propose different analyses of what counts as substandard. For AV, a person's cognitive and volitional abilities set the standard: a substandard failure is one that the agent had the ability and opportunity to avoid. An agent fails in a substandard way when she fails to perform according to her abilities. For QW, the standard is set by an appropriate quality of will: a substandard failure is one that the agent would have avoided if she had had an appropriate quality of will. An agent fails in a substandard way when she fails because of a problematic quality of will.

34 Johnson King, "Accidentally Doing the Right Thing"; Sliwa, "Moral Worth and Moral Knowledge."

## 5. CONCLUSION

A person cannot freely perform action *A* without satisfying some epistemic requirements. Plausibly, for example, to freely do *A*, a person must know that she is doing *A*. However, as I have shown, the epistemic condition on moral responsibility for *A* goes beyond the epistemic requirements for freely doing *A*. To be praiseworthy for a right action *A*, a person must know not just that she is doing *A* but also that *A* is morally right. And according to the strict view that only clear-eyed wrongdoings may be blameworthy, the epistemic condition on blameworthiness for a wrong action *A* requires not simply that the agent know that she is doing *A* but that she know that *A* is morally wrong. As we saw, some views adopt a looser epistemic condition according to which a person is blameworthy for a wrong action *A* only if she knows or should know that *A* is morally wrong. I examined two ways to analyze the locution “should know.” According to AV, a person should know that *A* is wrong just in case she has the ability and opportunity to form that knowledge. According to QW, a person should know that *A* is wrong just in case she would form that knowledge if she had an adequate quality of will. Although they propose slightly different analyses of “should know,” both AV and QW clearly impose an epistemic condition on moral responsibility. Challenges to this condition have thus proven unsuccessful.<sup>35</sup>

University of Oklahoma  
montminy@ou.edu

## REFERENCES

- Amaya, Santiago. “Slips.” *Noûs* 47, no. 3 (September 2013): 559–76.
- Anscombe, Elizabeth. *Intention*. Oxford: Blackwell, 1957.
- Arpaly, Nomy. *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford: Oxford University Press, 2003.
- Arpaly, Nomy, and Timothy Schroeder. *In Praise of Desire*. Oxford: Oxford University Press, 2014.
- Björnsson, Gunnar. “Explaining (Away) the Epistemic Condition on Moral Responsibility.” In Robichaud and Wieland, *Responsibility*, 146–62. Oxford: Oxford University Press, 2017.
- Clarke, Randolph. “Blameworthiness and Unwitting Omissions.” In *The Ethics*

35 I am grateful to an anonymous reviewer for this journal for helpful comments.

- and *Law of Omissions*, edited by Dana Kay Nelkin and Samuel C. Rickless, 63–83. Oxford: Oxford University Press, 2017.
- . “Ignorance, Revision, and Commonsense.” In Robichaud and Wieland, *Responsibility*, 233–51. Oxford: Oxford University Press, 2017.
- . *Omissions: Agency, Metaphysics, and Responsibility*. Oxford: Oxford University Press, 2014.
- Fischer, John Martin, and Mark Ravizza. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press, 1998.
- Ginet, Carl. “The Epistemic Requirements for Moral Responsibility.” *Philosophical Perspectives* 14, Action and Freedom (2000): 267–77.
- Harman, Elizabeth. “Does Moral Ignorance Exculpate?” *Ratio* 24, no. 4 (December 2011): 443–68.
- Johnson King, Zoë. “Accidentally Doing the Right Thing.” *Philosophy and Phenomenological Research* 100, no. 1 (January 2020): 186–206.
- Levy, Neil. *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. Oxford: Oxford University Press, 2011.
- Lewis, David. “Causation.” *Journal of Philosophy* 70, no. 17 (October 1973): 556–67.
- Markovits, Julia. “Acting for the Right Reasons.” *Philosophical Review* 119, no. 2 (April 2010): 201–42.
- McKenna, Michael. *Conversation and Responsibility*. Oxford: Oxford University Press, 2012.
- Mele, Alfred. “Moral Responsibility for Actions: Epistemic and Freedom Conditions.” *Philosophical Explorations* 13, no. 2 (2010): 101–11.
- Montminy, Martin. “Doing One’s Reasonable Best: What Moral Responsibility Requires.” *Journal of the American Philosophical Association* 2, no. 1 (Spring 2016): 55–73.
- Nelkin, Dana K. *Making Sense of Freedom and Responsibility*. Oxford: Oxford University Press, 2011.
- Raz, Joseph. “Being in the World.” In *From Normativity to Responsibility*, 227–54. Oxford: Oxford University Press, 2011.
- Robichaud, Philip, and Jan Willem Wieland, eds. *Responsibility: The Epistemic Condition*. Oxford: Oxford University Press, 2017.
- Rudy-Hiller, Fernando. “The Epistemic Condition for Moral Responsibility.” *Stanford Encyclopedia of Philosophy* (Fall 2018). <https://plato.stanford.edu/archives/fall2018/entries/moral-responsibility-epistemic/>.
- Sher, George. “Blame and Moral Ignorance.” In Robichaud and Wieland, *Responsibility*, 101–16. Oxford: Oxford University Press, 2017.



- . *Who Knew? Responsibility without Awareness*. Oxford: Oxford University Press, 2009.
- Sliwa, Paulina. “Moral Worth and Moral Knowledge.” *Philosophy and Phenomenological Research* 93, no. 2 (September 2016): 393–418.
- . “On Knowing What’s Right and Being Responsible for It.” In Robichaud and Wieland, *Responsibility*, 127–45. Oxford: Oxford University Press, 2017.
- Smith, Angela. “Responsibility for Attitudes: Activity and Passivity in Mental Life.” *Ethics* 115, no. 2 (January 2005): 236–71.
- Smith, Michael. *The Moral Problem*. Oxford: Blackwell, 1994.
- . “Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion.” In *Weakness of Will and Practical Irrationality*, edited by Sarah Stroud and Christine Tappolet, 17–38. Oxford: Clarendon Press, 2003.
- Strawson, Peter. “Freedom and Resentment.” *Proceedings of the British Academy* 48 (1962): 1–25.
- Talbert, Matthew. “Akrasia, Awareness, and Blameworthiness.” In Robichaud and Wieland, *Responsibility*, 14–63. Oxford: Oxford University Press, 2017.
- . “Omission and Attribution Error.” In *The Ethics and Law of Omissions*, edited by Dana Kay Nelkin and Samuel C. Rickless, 17–35. Oxford: Oxford University Press, 2017.
- Vargas, Manuel. *Building Better Beings: A Theory of Moral Responsibility*. Oxford: Oxford University Press, 2013.
- Wolf, Susan. *Freedom within Reason*. Oxford: Oxford University Press, 1990.

## KANT AND THE PROBLEM OF UNEQUAL ENFORCEMENT OF LAW

Daniel Koltonski

THE QUESTION of when disobeying the law may be justified as a means of protesting or otherwise resisting injustice is often narrowed from the start by the assumption that the only plausible candidate for such justified disobedience, at least in contemporary Western democracies, is civil disobedience.<sup>1</sup> There have been some recent exceptions, most notably Candice Delmas's series of arguments that standard liberal justifications for a duty to obey the law—fairness, justice, samaritanism, among others—will, in some circumstances, actually justify not only civil disobedience but also *uncivil* disobedience, including violent or destructive disobedience.<sup>2</sup> One figure to whom it might seem we cannot look for help justifying any such disobedience, particularly violent or destructive disobedience, is Kant, for in *The Doctrine of Right* and elsewhere he explicitly rejects not only revolution but also any resistance by citizens that aims to compel states to reform themselves. Indeed, it is not clear that Kant would even countenance what most liberals would take to be paradigmatic cases of justified civil disobedience.<sup>3</sup>

My aim here is to show that, in fact, the Kantian account of the legitimate state has the resources for a distinctive justification of principled disobedience,

- 1 The paradigm version of this view is Rawls's account in *A Theory of Justice*. Others have responded to Rawls's narrow account of justifiable disobedience by expanding the scope of what can count as civil (and so justifiable) disobedience. See, for instance, Brownlee, *Conscience and Conviction*.
- 2 These arguments are collected in her *A Duty to Resist*. Other recent exceptions are Shelby, *Dark Ghettos*; and Pasternak, "Political Rioting."
- 3 *The Doctrine of Right* is the first half of Kant's *The Metaphysics of Morals*. All citations to Kant are to the Akademie numbers listed in the margins of most editions; unless otherwise stated, all English translations are from *Practical Philosophy*, translated and edited by Mary Gregor. Kant does appear once in *The Doctrine of Right* to allow for individuals to disobey the law: "Obey the authority who has power over you (in whatever does not conflict with inner morality)" (*The Metaphysics of Morals*, 6:371). But he does not explain what he means by "inner morality," and the disobedience here seems more akin to conscientious objection than to disobedience as public protest or resistance.

including even violent or destructive disobedience, that applies to citizens of contemporary Western democracies. The argument is thus a contribution to this larger project of justifying on squarely liberal grounds a wider variety of ways citizens might resist injustice and try to compel reform. As Kant himself would reject this argument, my claim is not that Kant's own account allows for such principled disobedience; it is rather that a recognizably Kantian account—one that does not abandon the core Kantian commitments—can (and perhaps should) allow for it. In this way, the argument is also a contribution to the larger project, pursued by many over the last few decades, of reviving and defending the basic Kantian approach to questions of political authority and obligation while also moving beyond some of the more austere parts of Kant's own account.<sup>4</sup>

In *The Doctrine of Right*, Kant argues that a community of free and equal persons is possible only under the rule of law. His argument is distinctive in part because of the importance it places on the *equal enforcement* of law as required to secure persons' rights against invasion by their fellows. Individuals will have varying amounts of coercive power at their disposal to defend their rights themselves, and the result of this unequal distribution of coercive power is that absent state enforcement that assures them of the security of their rights, some will be vulnerable to domination by others, a condition incompatible with equal freedom. And so, for Kant, a civil condition—the condition under which citizens are governed by legitimate law, law they must obey and otherwise uphold—has two constitutive parts, the latter of which solves this “problem of assurance”: “What is to be recognized as belonging to [citizens] is determined by law and is allotted to [them] by adequate power.”<sup>5</sup>

What opens the door to the justifiability of disobedience that aims to compel a state to reform itself, even violent or destructive disobedience, is that a state may achieve the first part of a civil condition (“what is to be recognized as belonging to [citizens] is determined by law”) and yet at the same time *fail* to achieve the second (what the law says belongs to citizens is *not* “allotted to [them] by adequate power”) by failing to enforce the law equally. When this failure is a failure to provide some citizens the assurance of the security of their rights, the state has thereby failed to bring them entirely out of a state of nature and into a civil condition; they are instead in an incompletely civil condition,

4 Some examples are Korsgaard, “Taking the Law into Our Own Hands”; Hill, “A Kantian Perspective on Political Violence”; Holtman, “Revolution, Contradiction, and Kantian Citizenship”; Flikschuh, “Reason, Right, and Revolution”; Ripstein, *Force and Freedom*; Stilz, *Liberal Loyalty* and “Provisional Right and Non-State Peoples”; and Sinclair, “The Power of Public Positions.”

5 Kant, *The Metaphysics of Morals*, 6:312. As I understand them, these two parts are necessary to and jointly sufficient for a civil condition.

one that contains local illegitimacies. And so, were these citizens to employ tactics of resistance, even violent or destructive ones, with the aim of compelling the state to reform itself so that it provides them with equal assurance of the security of their rights, they would be using force to bring about (or at least get closer to) a full civil condition, and in that way, their use of force will count as furthering rather than undermining the project of governance by legitimate law.

Importantly, whether a Kantian account can justify resistance to injustice on these grounds is not merely a theoretical question, for many existing states fail in significant ways to provide their citizens with equal assurance of the security of their rights. One clear case, as I will argue, is that of wage and hour laws in the United States. While American law defines for workers their wage rights, the United States fails to enforce those laws against employers and so fails to provide those workers with assurance of the security of those rights within the employment relationship. Whether workers get the wages to which they are legally entitled is, in the end, up to employers, and so, unsurprisingly, wage theft in the United States is widespread. Thus, within the employment relationship, workers and employers are not yet fully in a civil condition. On the Kantian argument I defend here, workers are permitted to use tactics of resistance that aim to compel those reforms to the United States' enforcement regime that are necessary for the security of their wage rights.

#### 1. EQUAL FREEDOM AND THE PROBLEM OF ASSURANCE

Kant argues that the private enforcement of rights is incompatible with equal freedom. We can capture the basic problem by thinking about how inequalities of coercive power among persons would affect the prospects for private enforcement of one's rights against others. Imagine a case where the relevant principles clearly define who owns what. Suppose, somewhat fancifully, that according to some valid principle of acquisition (finders keepers, say), one clamshell on the beach is mine and another is yours.<sup>6</sup> Because on Kant's account a right to a thing includes the authorization to coerce, that we each have a right to our respective shells would mean, in part, that we each may use coercion against the other should that other try to take our shell. The problem is that in this situation you may have available more power to coerce me than I have to coerce you (or vice versa). If you do, your greater coercive power is a powerful incentive for me to respect your claim, but at the same time, your greater power means that you lack a similar incentive to respect my claim. You are thus able to secure your shell for yourself and against me while I am unable similarly to secure my shell for my-

6 I have borrowed this example from Sinclair, "The Power of Public Positions," 32.

self and against you. And so, were we both authorized by right to use coercion to defend our rights ourselves, this inequality of coercive power would make it such that only you can take advantage of this authorization: you can oblige me to respect your right to your shell in a way that I cannot in turn oblige you.

That your greater coercive power here gives you assurance of the security of your right is not itself a problem, as this assurance is partly what makes you independent of me: because you can defend your right against invasion by me, your enjoyment of your right does not depend on my decision to invade it or not, and in that way, it is independent of my will. The problem is that because of your greater coercive power, I lack this same assurance of the security of my right: you could decide to invade my right simply because you wanted my shell for yourself and I would be unable to stop you. My property—my shell—is thus available to you, just as your property is, as a means for you to use for your own private ends, and so I depend on your will for my enjoyment of this right. Were the principle of acquisition above authoritative in these conditions of unequal coercive power—were we to have conclusive rights to our shells—I would then be required to respect your right even though, because of my dependence on you, I would lack assurance that you will respect my right. As Thomas Sinclair puts it, I would be required by right “to leave [myself] open to be taken advantage of” by you.<sup>7</sup> But because such a requirement is incompatible with our equal freedom, the principle is not authoritative and we do not have conclusive rights to our respective shells.

A distinctive feature of Kant’s account is thus that our unequal access to coercive power means that we cannot settle what our rights are simply by appeal to what the relevant principles say, for while the principles may assign to each of us ownership of our respective shells, our unequal access to the coercive power we require to secure our rights ourselves undermines our equal freedom. In this way, the freedom at stake is not merely formal freedom: it is not enough that a valid principle of right gives you ownership of the shell and, in doing so, authorizes the use of coercion to protect this right; it matters also whether you actually have access to the coercive power—and, if so, how much relative to others’—needed actually to secure your rights. You are not in fact free if you are able to enjoy the rights that constitute your freedom only at the whim of some other person(s).

Now, to be precise, what matters for the question of assurance is not inequalities in coercive power but rather inequalities in the coercive threat persons pose to one another. Presenting it in terms of inequalities of coercive power, as in the clamshell case, is simply a way to make this problem clear and vivid, for one standard way for you to pose more of a coercive threat to me in some situation

7 Sinclair, “The Power of Public Positions,” 33.

than I do to you is for you to have more coercive power at your disposal than I do.<sup>8</sup> For instance, that you have a gun (and I do not) will often result in an inequality of coercive threat between us such that you are able to defend your right to your shell. But not always. That you have a gun will pose much less of a threat to me, for instance, if you have fallen asleep and I can take your shell without waking you up. What matters, at bottom, is the incentive your coercive power actually gives me to respect your right and, in particular, whether the threat that power poses is able to oblige me to respect your right. It may fail to oblige me either because your gun has jammed (you have lost that coercive power and so any threat it might pose) or because you have fallen asleep (this coercive power no longer translates into a threat). That the issue is ultimately about inequalities of coercive threat matters because it is exceedingly unlikely that any particular person, regardless of how much coercive power they have at their disposal, will be able to oblige others consistently across time and changing circumstances to respect their rights. The inequalities of coercive threat between persons will thus vary: your enjoyment of your rights will, at times, be at my whim and my enjoyment of mine will, at times, be at your whim, and so we will each be threats to the security of the other's rights. As a result, neither of us will be able to secure our rights by means of private enforcement.

This problem of assurance features in Kant's core argument for the state. The argument begins with a state of nature occupied by idealized human beings—they are entirely "good and right-loving"—and this is so that, as Arthur Ripstein observes, the case for the state will not depend on any "empirical defects of the state of nature, such as self-preference and limited knowledge."<sup>9</sup> According to this argument, even these idealized human beings will require a state:

However [good and right-loving] human beings might be, it still lies a priori in the rational idea of such a condition (one that is not rightful) that before a public lawful condition is established, individual human beings ... can never be secure against violence from one another, since each has its own right to do *what seems right and good to it* and not to be dependent upon another's opinion about this.<sup>10</sup>

In this state of nature, everyone is subject to the rules of right defining their equal freedom. What they are not subject to is an authority's judgments of those

8 I am indebted to an anonymous reviewer for pressing me about these issues.

9 Ripstein, *Force and Freedom*, 146.

10 Kant, *The Metaphysics of Morals*, 6:312. The bracketed text comes from Arthur Ripstein's translation in *Force and Freedom*, 146. The Gregor translation has it as "well-disposed and law abiding."

rules and how they apply; they are each instead self-governing, acting on their own judgments. Kant's claim here is that in a situation where even entirely good and right-loving human beings govern themselves according to their own judgments of right, they cannot be secure from violence. This insecurity stems from the absence of a mechanism for resolving their inevitable disagreements about right: when such disagreements arise, each will act on their own judgment of the matter, standing up to others in defense of what they take right to be. From their own perspective, their use of coercion will be authorized by right, but because of their disagreements about right, their uses of coercion will oppose one another. Everyone will thus be subject to coercive threat—doing only what you judge consistent with right will be no protection from others who may disagree—and, in that way, no one will be secure from violence. In the state of nature, a person will lack assurance that others will respect her rights (as she sees them), and as a result, she will not be required to respect their rights (as she sees them).

The solution, Kant argues, is a state that not only settles by legislation and adjudication what counts as right among them but also *enforces* those laws and legal decisions so as to secure for each their rights so defined:

So, unless [a human being] wants to renounce any concepts of right, the first thing it has to resolve upon is the principle that it must leave the state of nature, in which each follows its own judgment, unite itself with all others (with which it cannot avoid interacting), subject itself to a public lawful external coercion, and so enter into a condition in which what is to be recognized as belonging to it is determined by *law* and is allotted to it by adequate *power* (not its own but an external power); that is, it ought above all else to enter a civil condition.<sup>11</sup>

The problem, however, is that Kant's argument here does not actually establish that if these idealized human beings are to be assured of the security of their rights, the state must operate as a central enforcement agent. In this state of nature, recall, the reason why they lack equal assurance is that rights are in dispute: it is because others disagree with her about right that she cannot be assured that they will respect her rights (as she sees them). This suggests that were they to solve the problem of disagreement, they would thereby also solve the problem of assurance.<sup>12</sup> Suppose they have solved the former problem with a state that is solely legislative and adjudicative. Provided that it is common knowledge

11 Kant, *The Metaphysics of Morals*, 6:312.

12 I depart here from the interpretation offered by Ripstein. He argues that for our idealized persons in the state of nature, the problem of assurance would persist even if the problem of disagreement were solved. See Ripstein, *Force and Freedom*, 146, 159.

among them—and there is no reason to think that it will not be—the fact that they all are entirely good and right-loving would seem to provide each of them with the equal assurance necessary for them to be required to respect the rights of others (as the state defines them), for what they are assured of is that those others will respect their rights (as the state defines them).

If the state must also be the central enforcement agent even among entirely good and right-loving human beings, it cannot be because equal assurance requires it. Perhaps it is instead the unequal dependence itself, apart from whether it results in unequal assurance, that is incompatible with equal freedom.<sup>13</sup> On this line of thought, even though I have assurance that you, entirely good and right-loving, recognize an obligation to respect my right to my shell, I still depend on your recognition of this in a way that because you pose more of a threat to me than I do to you, you do not depend on my recognition of the corresponding obligation to you. Your greater threat gives me an additional incentive to respect your right—it obliges me to do so—and this is an incentive that you do not have to respect mine. Even though we enjoy equal assurance—your ability to oblige me here is superfluous and my inability to oblige you does not make my rights any less secure than yours—this inequality of threat is still incompatible with equal freedom, for it is nevertheless true that the security of your right in this instance is independent of my will in a way that the security of mine is not independent of yours. Among the entirely good and right-loving, then, what the state as a central enforcement agent does is to make the security of everyone's rights independent of the wills of others, and equally so. It deploys its coercive power so as to oblige everyone equally to respect the rights of everyone else, and even though no one requires this incentive to respect others' rights, that the state provides it to all is necessary for equal freedom.

One might reasonably wonder whether this argument succeeds as a justification of the enforcement functions of the state when persons are idealized in this way. If the inequalities of coercive threat among them do not undermine the security of their rights—the fact that they are all entirely good and right-loving secures them, and equally so—then, one might think, those inequalities do not undermine their equal freedom. Fortunately, whatever one thinks of this justification, Kant provides a noticeably more powerful justification for why realizing equal freedom among human beings like us, ones who are decidedly *not* entirely good and right-loving, requires the state to be the central enforcement agent.

13 Sinclair's interpretation does not make clear whether the problem at issue is unequal dependence itself or rather the insecurity of one's rights that can result from such unequal dependence. See Sinclair, "The Power of Public Positions," 33. This is also true of Anna Stilz's account. See Stilz, *Liberal Loyalty*, 51.



In a different passage, Kant does not idealize the human beings in the state of nature, and in this version of the argument, the problem of assurance reemerges:

No one is bound to refrain from encroaching on what another possesses if the other gives him no equal assurance that he will observe the same restraint toward him. No one, therefore, need wait until he has learned by bitter experience of the other's contrary disposition; for what should bind him to wait till he has suffered a loss before he becomes prudent, when he can quite well perceive within himself the inclination of human beings generally to lord it over others as their master (not to respect the superiority of the rights of others when they feel superior to them in strength and cunning)?<sup>14</sup>

In this state of nature, human beings are not entirely good and right-loving but instead possess "the inclination ... to lord it over others as their master." Here, then, though you and I may agree about what right requires in some situation, that you have, and perceive yourself to have, more threat power than I do in that situation means that though you are assured of the security of your right, I cannot be similarly assured of the security of mine: you might act on this inclination to dominate me and invade what you yourself acknowledge to be my rights, and were you to do so, I would be unable to stop you. My dependence on you here makes my rights insecure. We thus require a state as a central enforcement agent to oblige us both, and equally so, to respect the other's rights, for it is only then that our rights will be equally secure and we equally free.

The question now concerns what the state as central enforcement agent must do to achieve the condition under which the rights of people like us are, as Kant puts it, "allotted to [them] by adequate *power*." On Ripstein's reading, the incentive that the state's enforcement power provides us to respect one another's rights has "two dimensions":

First, it assures the private right holder that the right will remain intact, even if another violates it. Second, it makes rights violations prospectively pointless. If a right holder is assured of a remedy, others will not normally have any incentive to violate rights, because a violator will expect to gain nothing and could possibly lose something through a violation.<sup>15</sup>

This might seem relatively straightforward. But it cannot be, as Ripstein has it, that what the state is to do is to assure "the private right holder that the right will remain intact," thereby making rights violations "prospectively pointless," for as

14 Kant, *The Metaphysics of Morals*, 6:307.

15 Ripstein, *Force and Freedom*, 165–66.

Sinclair points out, that would require that the state's enforcement apparatus be fully effective, remedying every rights violation in society, which is an impossible standard to satisfy.<sup>16</sup> The problem seems to me, however, to be even worse. Assuring a remedy for every rights violation would not make those violations prospectively pointless, for if people are generally inclined "to lord it over others as their master," the rights violation itself, as an assertion of power, will often be the point, not some further payoff. To render these violations prospectively pointless, as Ripstein's reading requires, the state would need to be such that it credibly promises not to remedy every rights violation after the fact but rather to block every one before it occurs.

So what then are the conditions the state must provide citizens such that their rights are allotted to them by adequate power? Sinclair argues that they are the conditions under which "I am not *more* assured of compliance with my claims of acquired right than you are of yours. For only that makes for a problem of *unequal* freedom."<sup>17</sup> On this reading, it is not that a citizen must be guaranteed that her rights are totally secure from violation by others but rather that she must have assurance of their security equal to that of others. Sinclair's version of this reading goes awry, however, in holding that equal assurance is secured when it is the case that "even if some are better positioned than others to violate acquired rights, they don't have a better prospect than others of maintaining the resultant configuration of external freedom."<sup>18</sup> What matters for equal assurance is not, as Sinclair has it, persons' chances relative to others at getting away with an unremedied rights violation but rather their chances relative to others at suffering one. That several people might suffer such a violation at the hands of the same perpetrator—someone better positioned than others to get away with rights violations—will not by itself mean that they do not enjoy equal assurance, for it may still have been that their chances at suffering a violation were roughly equal.<sup>19</sup> The problem arises instead when some person is particularly vulnerable, more so than others, to such violations, whether at the hands of one perpetrator or several.

On my reading, then, a state secures citizens' rights with adequate power when it has more coercive power than any particular citizen across a full range of situations and circumstances—it is able to oblige them to respect others' rights—and when it wields that power in those situations such that the security it provides to the rights of those involved is distributed equally among them so that no one is specially vulnerable to an unremedied rights violation. This does

16 Sinclair, "The Power of Public Positions," 39.

17 Sinclair, "The Power of Public Positions," 46.

18 Sinclair, "The Power of Public Positions," 46.

19 I am indebted to an anonymous reviewer for pressing me about Sinclair's account.

leave open the possibility that sometimes one citizen will suffer a rights violation that is not remedied—the assurance the state provides is not total—but so long as that citizen's prospects at suffering such a violation are not meaningfully worse than others' prospects, that possibility is not incompatible with equal freedom, for she is no more dependent on others for the security of her rights than her fellows are. And, unlike Ripstein's reading, this reading has the advantage that the standard it sets is one a state *can* meet. (The Kantian account will succeed only if the state can do what the account claims the state must do.)

The literature on Kant's assurance argument is almost wholly concerned, first, with defending the claim that inequalities of coercive power (or threat) are incompatible with equal freedom and, second, with showing that the state as the central enforcement agent can solve this problem of assurance. What has not gotten attention is the fact that this requirement that the state provide citizens with equal, though not total, assurance is still a *very* demanding one. Even if it is possible for a state to solve the problem of assurance, actually doing so will be quite difficult. It is far from obvious, for instance, that existing states even come close to succeeding at this task.<sup>20</sup> And since providing equal assurance is, on Kant's account, a core function of the state—indeed, one on which the state's claim to legitimate authority rests—it is an important question what the implications are for a state's authority when it does not in fact succeed.

This question is not a merely theoretical one. Many existing states, for instance, contain large inequalities of economic power, inequalities that tend to generate corresponding inequalities of political power and influence; these inequalities of political power and influence tend, in turn, to shape both the decisions lawmakers make about funding enforcement agencies and the decisions officials within those agencies make about enforcing the law. As a result, these states tend to enforce the law more stringently when poor citizens are the violators or wealthy ones the victims than when wealthy citizens are the violators or poor ones the victims. Across a range of rights, wealthy citizens will tend to be more assured of the security of their rights than poor citizens will be of the security of theirs. Indeed, the police and courts may even come to function as tools available for the wealthy and powerful to use not only to protect their rights but also to violate the rights of others with relative impunity. One clear example of this phenomenon is the severe and widespread underenforcement of wage and hour laws in the United States. As the next section explains, in this area of law, the United States not only fails to provide one group of Americans with equal

20 Sinclair seems to admit as much: "I am not claiming that any existing state secures the condition of equal assurance. But it is not implausible to think that a realistically attainable state might" ("The Power of Public Positions," 47).

assurance of the security of their rights but also, arguably, does not even try to do so.

## 2. WAGE THEFT IN THE UNITED STATES

In the United States, wage theft—“the failure to pay workers the full wages to which they are legally entitled”—is severely underpoliced and, as a result, is fairly widespread.<sup>21</sup> Wage theft comes in many forms: underreporting hours worked, taking illegal deductions from paychecks, or failing to pay the legally mandated overtime rate for hours worked over forty hours per week.<sup>22</sup> But according to Daniel Galvin, the “most pernicious type of wage theft is minimum wage non-compliance,” for violations of minimum wage laws “disproportionately affect the most vulnerable workers in society: immigrants, people of color, less-educated workers, younger workers, women, and low-wage workers who can least afford to be underpaid.”<sup>23</sup> Wage theft in the United States is a clear and compelling example of a state’s failure to provide certain of its citizens—here, some of its least powerful—equal assurance of the security of their legally defined rights.

The losses from wage theft in the United States are quite large, both absolutely and as a proportion of income. A study conducted in 2008 of the losses to workers in three American cities (New York, Los Angeles, and Chicago) from all forms of wage theft combined found the following:

More than two-thirds (68 percent) of our sample experienced at least one pay-related violation in the previous work week. The average worker lost \$51, out of average weekly earnings of \$339. Assuming a full-time, full-year work schedule, we estimate that these workers lost an average of \$2,634 annually due to workplace violations, out of total earnings of \$17,616. That translates into wage theft of 15 percent of earnings.<sup>24</sup>

Another report, this one focusing only on minimum wage violations in the ten largest states and using data from 2013 to 2015, found that in these states, “2.4 million workers lose \$8 billion annually (an average of \$3,300 per year for year-round workers) to minimum wage violations—nearly a quarter of their earned wages.”<sup>25</sup> If this report’s findings hold true also for workers in the remaining forty

21 Cooper and Kroeger, “Employers Steal Billions from Workers’ Paychecks Each Year,” 4.

22 Cooper and Kroeger, “Employers Steal Billions from Workers’ Paychecks Each Year,” 4. See also Bernhardt et al., “Broken Laws, Unprotected Workers.”

23 Galvin, “Deterring Wage Theft,” 325.

24 Bernhardt et al., “Broken Laws, Unprotected Workers,” 5.

25 Cooper and Kroeger, “Employers Steal Billions from Workers’ Paychecks Each Year,” 1.

states, upward of \$15 billion is stolen from workers in the United States via minimum wage violations each year.<sup>26</sup>

In his study concerning the prospects for more effective enforcement of wage and hour laws, Galvin offers these further statistics about minimum wage violations:

An estimated 16.9 percent of low-wage workers experienced a minimum wage violation in 2013. Those workers worked on average 32 hours per week and earned an hourly wage of \$5.92. Had they earned their state's minimum wage, they would have earned, on average, an hourly wage of \$7.62, which means they lost 23 percent of their income (\$1.76 per hour).<sup>27</sup>

It is admittedly quite difficult to gather data about wage violations, particularly in low-wage sectors, and as a result, the estimates of the extent of wage theft vary quite a bit. But as Galvin notes, "while an estimated income loss of 23 percent may seem high, it is actually toward the lower end of other published estimates."<sup>28</sup> We can thus confine ourselves to the lower-end estimates and still the problem of wage theft is quite large: millions of workers in the United States are victims of it, many of them routinely so, and the thefts are significant, both in the absolute amount and as a percentage of their income. Indeed, if we use the \$15 billion estimate, the money stolen from workers via minimum wage violations alone is more than the value of what is stolen in property crimes in the United States: "According to the FBI, the total value of all robberies, burglaries, larceny, and motor vehicle theft in the United States in 2015 was \$12.7 billion."<sup>29</sup>

Why might wage theft be such a pervasive problem in the United States? One important reason seems to be that government, both the federal government and many state governments, does very little to enforce the wage and hour laws on the books. For instance, over the past seventy years, the number of federal investigators in the Wage and Hour Division of the Department of Labor has remained static even though the workforce is now six times larger, and the result is that employers are exceedingly unlikely to be investigated (only a 0.5 percent chance in 2012).<sup>30</sup> Indeed, between 1980 and 2015, the number of cases inves-

26 Cooper and Kroeger, "Employers Steal Billions from Workers' Paychecks Each Year," 28.

27 Galvin, "Deterring Wage Theft," 330–31.

28 Galvin, "Deterring Wage Theft," 331. As he notes, a 2014 study commissioned by the US Department of Labor found that in 2011 the average income loss to minimum wage violations in New York and California was 37 percent and 49 percent, respectively. See US Department of Labor, *The Social and Economic Effects of Wage Violations*.

29 Cooper and Kroeger, "Employers Steal Billions from Workers' Paychecks Each Year," 28.

30 Galvin, "Deterring Wage Theft," 327.

tigated by the Wage and Hour Division has shrunk by 65 percent.<sup>31</sup> In light of this, a worker might reasonably decide to pursue a wage violation claim against her employer herself. Were she to do so, she would find that in many jurisdictions, the process for filing such a claim is quite difficult, seemingly arbitrarily so. For instance, in Iowa, the complaint process “contains a multitude of procedural obstacles that may actively discourage workers from pursuing claims.”<sup>32</sup> Class action lawsuits may also be an option, though normally only against large companies and only for those employees who have not already agreed, as a condition of employment, to take any employment disputes to binding arbitration.<sup>33</sup>

The results are unsurprising. As Galvin observes, “because the probability of detection in the United States is so low, the literature on minimum wage compliance has long concluded that in actuality government-imposed penalties do not seriously affect the employer’s incentives.”<sup>34</sup> Indeed, even in the rare cases where employers are investigated and found to have committed wage theft, the settlements reached often do not require that they pay all the back wages owed but rather only up to two years’ worth. As Orley Ashenfelter and Robert S. Smith point out, “the requirement that a violating employer merely pay to em-

31 Cooper and Kroeger, “Employers Steal Billions from Workers’ Paychecks Each Year,” 5.

32 Gordon et al., *Wage Theft in Iowa*, 14. They add:

Under existing procedural rules, workers with limited literacy skills, limited English, or those who simply lack time or access to documentation, or have no permanent mailing address, would all effectively be precluded from filing claims. And under current rules, where a worker’s burden to provide additional documentation per-  
versely increases in proportion to an employer’s nonresponsiveness, even workers who manage to file initial written claims are at high risk of having their cases “closed” at any point in the process for a myriad of procedural reasons. (14)

Additionally, workers are on their own during the process:

Though employers are allowed to have attorneys or other third parties represent them in the claims process, workers are not. In fact, IWD [Iowa Workforce Development] will close a worker’s case if it learns that an attorney or other third party (a pastor, union representative, or community organizer, for example) is assisting the worker in contacting the employer, communicating with enforcement agencies, or using other means to try [to] recover the worker’s wages. (14)

33 As Gilman explains, employees must typically pursue a wage claim in arbitration individually, which raises the costs to the employee: “To begin with, employees need a lawyer if they want any hope of prevailing—although this is nearly impossible in the absence of a class action option. In addition, arbitrators charge hundreds of dollars an hour, and many agreements require employees to pay part of the costs” (“Supreme Court Ruling against Class Action Lawsuits Is a Blow for Workers—and #MeToo”). For a thorough discussion of the use of binding arbitration in employment disputes, see Colvin, “Mandatory Arbitration and Inequality of Justice in Employment.”

34 Galvin, “Deterring Wage Theft,” 328.

ployees a fraction of the difference between the minimum and the actual wage received does not constitute a penalty for noncompliance at all.”<sup>35</sup> Furthermore, as Galvin notes, “civil or criminal penalties are rare, reserved for cases of employer retaliation, repeat, or ‘willful’ violations.”<sup>36</sup> From the standpoint of American employers, then, it is as if there is no state enforcement of wage and hour laws at all.<sup>37</sup>

The legal regime of property rights in the United States offers little to no assurance to workers that their rights to their wages will be respected. These workers are thus dependent on their employers in a way incompatible with their equal freedom, for given the inequality of power generally between employers and workers, whether their rights to their wages are respected depends on the will of their employer. Of course, particular employers may have various other incentives either to respect these rights fully or at least not to violate them egregiously (e.g., concerns about productivity effects or bad PR, or perhaps concern for employee well-being or even moral principle) but what is missing is the incentive required, on the Kantian account, to solve the problem of assurance, the incentive provided equally to all by the state’s promise of enforcement.

### 3. WAGE THEFT AND PRINCIPLED DISOBEDIENCE

As we saw, for Kant, the solution to the problems of the state of nature is a civil condition, and this condition has two constitutive parts: “What is to be recognized as belonging to [citizens] is determined *by law* and is allotted to [them] by adequate *power*.”<sup>38</sup> What is troubling about the situation of low-wage workers in the United States is that they are not entirely in such a civil condition with employers, for while the law does define their wage rights—and so “what is to be recognized as belonging to [them] is determined *by law*”—it is not the case that their wages are “allotted to [them] by adequate *power*.” Though Kant nowhere explicitly discusses the possibility of persons being in such a situation, we can ask what a Kantian account should say about their rights and obligations in it.<sup>39</sup>

35 Ashenfelter and Smith, “Compliance with the Minimum Wage Law,” 337, quoted in Galvin, “Deterring Wage Theft,” 328.

36 Galvin, “Deterring Wage Theft,” 328.

37 This problem of wage theft admittedly also involves the failure of adjudicative and administrative agencies to apply the law, and so it seems also to impact the first of Kant’s two parts of a civil condition. As my interest here is to see how far one can get with just a focus on failures of enforcement, I am bracketing this other important aspect of the wage-theft case. I am indebted to an anonymous reviewer for pressing me on this point.

38 Kant, *The Metaphysics of Morals*, 6:312.

39 Kant does discuss a situation like this when it comes to states, for he argues, in *Toward Per-*

Since the state is failing to allot to these workers their wages by adequate power, what may they do themselves to secure their legal rights to their wages? May they, at most, raise objections and issue calls for reform, with the implication that while they await those reforms, they must confine themselves to the wholly inadequate avenues of redress the legal system makes available to them as victims of wage theft? Or may they engage in acts of disobedience and resistance, perhaps even violent or destructive acts, with the aim of compelling the state to reform itself and improve its enforcement of wage laws?

The answer Kant himself would give is clear, as he argues that even if a constitution is “afflicted with great defects and gross faults and [is] in need eventually of important improvements,” citizens must nevertheless obey its laws, for “the defects attached to it must instead be gradually removed by reforms the state itself carries out.”<sup>40</sup> The state’s duties to reform itself are not ones citizens may coerce it to fulfill, for instance, by threatening resistance or revolution; the most they may do is raise complaints and objections.<sup>41</sup> Kant’s answer, then, would be that even those left vulnerable to wage theft by the state’s failure to enforce wage and hour laws must nevertheless obey the law. They must put up with this great defect in the legal system and confine themselves to whatever inadequate avenues of redress the legal system makes available to victims of wage theft. They may protest the legal system’s failure to protect their wage rights by raising objections for officials to consider, but on Kant’s view, that seems to be all they may do to push for reform.<sup>42</sup> Even civil disobedience is ruled out.

My aim here is to show that a different answer is possible, one that is arguably truer to the basic commitments of the Kantian account. These workers may engage in acts of disobedience and resistance—acts of strategic illegality—to compel the United States to reform its enforcement of wage laws so that it provides them the equal assurance of the security of their wage rights that is partly constitutive of a civil condition. This resistance will be what Delmas calls “principled disobedience”: “politically or morally motivated resorts to illegality in the

---

*petual Peace* and elsewhere, against a transition to a world state and in favor of the establishment of a federation of states (8:357). The difference is that when it comes to states, such a situation is one that, on Kant’s view, they ought to strive for, while when it comes to persons, it seems one that they would need to overcome. I am indebted to an anonymous reviewer for pointing me to Kant’s discussion of a federation of states.

40 Kant, *The Metaphysics of Morals*, 6:372; see also 6:322.

41 Kant, *The Metaphysics of Morals*, 6:319. Kant claims that “a people has a duty to put up with even what is held to be an unbearable abuse of supreme authority” (6:320).

42 Kant says that even when it comes to an action done by the executive power (“the ruler”) that is contrary to law, citizens “may indeed oppose this injustice by complaints but not by resistance” (*The Metaphysics of Morals*, 6:319).



opposition or refusal to conform to the system's dominant norms."<sup>43</sup> And since such disobedience may involve property damage or destruction, it will not fall under the umbrella of civil disobedience. As civil disobedience is often the only kind of principled disobedience liberals are comfortable admitting can be justified, the claim here is that the Kantian account has the resources to be somewhat more radical than the standard liberal view.

### 3.1. *The Possibility of Local Illegitimacy*

Kant's claim about citizens' duties to obey the law, even in the face of "great defects," is based on the ideal of the fully rightful state, an ideal that, as he puts it, "serves as a norm for every actual union into a commonwealth ([and] hence serves as a norm for its internal constitution)."<sup>44</sup> As Anna Stilz explains, this ideal acts as a guide for reforming the current defective condition:

Even if their . . . existing constitution falls grossly short of this norm, [this constitution] is still legitimately binding because the existing juridical union is an important prerequisite for progress *toward* a republican constitution. It is for the ruler to reform the constitution legally, not for the subject to use force to usher in a new order of things.<sup>45</sup>

Despite its failure to enforce wage and hour laws, workers must still recognize the American state's legitimacy, for its legitimacy is a prerequisite for the reforms that would secure their wage rights. As Stilz's gloss makes clear, Kant's position assimilates any use of force by citizens against the state, even force exercised to compel the state to take their objections seriously, with resistance that rises to the level of revolution ("usher[ing] in a new order of things").

Katrin Flikschuh makes a similar point about the defective state: "Kant regards entrance into the civil condition as a necessary but not a sufficient condition of possible relations of Right. The fact that a regime fails to meet the sufficiency condition does not entitle the people to violate the necessity condition: one does not ensure justice by challenging legitimacy."<sup>46</sup> The result, as Flikschuh makes clear, is that "for Kant, most existing states are legitimate yet more or less unjust."<sup>47</sup> This distinction between justice and legitimacy is indeed crucial, and my argument here does not deny it, for its claim is not merely that the state's failure to enforce wage laws is unjust; it is rather that it is a failure to bring workers

43 Delmas, *A Duty to Resist*, 42.

44 Kant, *The Metaphysics of Morals*, 6:313.

45 Stilz, "Provisional Right and Non-State Peoples," 216.

46 Flikschuh, "Sidestepping Morality," 138.

47 Flikschuh, "Sidestepping Morality," 138.

and their employers entirely into a civil condition in the first place, and thus that in this context, the state is not fully legitimate. Some great defects may not pose a challenge to legitimacy, but the failure to enforce workers' wage rights is one that does.

The challenge to legitimacy at issue here is, as it were, a *local* one. The claim is thus that a state can be generally legitimate while also locally illegitimate. Why might this be? In Kant's core argument for the state, you must enter into a civil condition with those with whom you cannot avoid interacting. A civil condition, as Kant has it, is relational: when you are in a civil condition, you are in it *with* particular others. So, being in a civil condition with some other means that in your interactions with them, what belongs to you both—your rights in those interactions—is determined by law *and* allotted to you both by adequate power. It is possible, of course, to be in a civil condition with some people and not with others. But it also seems possible to be in a civil condition with another person not generally but only with regard to some interactions, for whether the state allots to us our rights with adequate power can vary across types of interactions in a society: in some interactions, the state may enforce the various rights involved equally while, in others, it may not. The state's failure in these latter interactions to allot some citizens their rights with adequate power would amount to a local illegitimacy, as one necessary constituent of a civil condition is absent. And the case of wage theft in the United States seems to me a clear instance of this. The American economy is structured in such a way that a large portion of the population cannot help but work for others for a wage, and so their unavoidable interactions with those others will involve the rights and obligations that the law sets out as constituting the employment relationship. But the legal system fails to enforce workers' wage rights, while it does enforce employers' rights, and so in their interactions with employers within this legally defined employment relationship, workers are not in a civil condition with employers. This relationship is instead the site of a local illegitimacy.

The claim here is thus not that the American state, as a whole, is illegitimate. When the failure to provide equal assurance is in this way local, the resulting illegitimacy will be as well. Allowing for such local illegitimacies is compatible with the basic thrust of the Kantian position, as presented by Stiliz and Flikschuh: the existing state, though grossly defective in particular ways and so locally illegitimate, nevertheless goes some distance toward bringing everyone into a civil condition, and because it at least does this, citizens are to respect the existing institution. Since American law, including wage and hour laws, does define a system of property rights that applies to all—the first part of a civil condition—anyone committed to entering into a civil condition with others will have no

reason not to recognize the existing legal system as achieving at least one major step toward such a condition. They are thus not to challenge the legitimacy the American state does possess by trying to “usher in a new order of things” via revolt or revolution. But they may aim to reform it so as to remove the defect and with it the local illegitimacy. And because the defect at issue is one of legitimacy, they may use force—and in doing so, disobey the laws—to bring about the necessary reforms so that the American state moves closer to achieving the conditions of fully legitimate rule across all citizens.

Indeed, the Kantian account ought to allow for this kind of local illegitimacy if it is to stay true to its claims about the importance of equal assurance to equal freedom. The American state does provide assurance to employers of their rights in this employment relationship—its coercive power obliges workers to respect those rights—so, on the assurance argument, employers are bound by wage and hour laws to pay workers what those laws declare them entitled to. But the state does not provide such assurance to the workers of their rights in this relationship, and this failure puts workers in exactly the position that, on Kant’s assurance argument, undermines claims of right in the state of nature. Consider the claim that despite the state’s failure to enforce wage rights, workers are bound to abide by the laws governing the employment relationship and so to respect employers’ rights. This amounts to claiming that they are required by right to leave themselves open to being taken advantage of by employers, a requirement that the assurance argument claims is incompatible with their equal freedom. As a result, they cannot be so required. The Kantian position, then, should be that these citizens need not confine themselves to whatever inadequate avenues of redress the legal system makes available to them as victims of wage theft; instead, they may use the coercive power available to them to try to compel reform.

In addition, this failure to provide workers with equal assurance of their wage rights changes the significance of the assurance of rights the state provides to employers. Because the state secures employers’ rights against workers (were an employee to attempt to steal from her employer, the state would be very likely to remedy it and punish her) but not vice versa, workers are obliged by the state’s power to respect employers’ rights even though employers are not obliged to respect theirs. Employers are thus not only left free by the state to engage in wage theft but also protected by the state from consequences, for the state will protect the proceeds of wage theft from any private remedies workers might attempt as if those proceeds were employers’ rightful property. The state’s selective enforcement of rights within this relationship thus functions as the exercise not of properly public coercive power but of private power that both reinforces and widens the power inequality between employers and workers in the employ-

ment relationship, making workers' rights to their wages even more vulnerable to invasion and so even less secure. In this context, the legal system's requirement that workers leave the defense of their wage rights to the state cannot help but be illegitimate and so not binding.

But why is it that workers not only may disobey the law here but also may use force to try to compel reform? The answer has to do with the status of the power that is used against them in this context. When employers engage in wage theft, taking advantage of this opportunity to steal wages with impunity, they reveal themselves to be individuals who are "not willing to submit" to a civil condition with workers.<sup>48</sup> The same is true, say, when employers hire lobbyists or form industry groups with the aim of undermining state- or federal-level efforts at reform. Exercising power in these ways against workers amounts to refusing to do what Kant's core argument for the state claims is "the first thing" right requires of them, which is to enter into a civil condition with those "with [whom they] cannot avoid interacting."<sup>49</sup> And the state, by failing to enforce workers' rights while also standing ready, under the guise of protecting employers' rights, to block workers from pursuing private avenues of redress, reveals itself to be an accomplice to employers' resistance to a civil condition. In a state of nature, resistance to entering into a civil condition amounts, on Kant's view, to a repudiation of right altogether, so one may use force to counter this sort of resistance and bring about a civil condition: "Each may impel the other by force to leave this state [of nature] and enter into a rightful condition."<sup>50</sup> What workers face here regarding their wage rights is simply a local version of this resistance, so they too may use force in response.

### 3.2. *Wage Theft and Principled Disobedience*

Because the failure to enforce wage laws is a defect of legitimacy, workers' acts of principled disobedience can count as efforts to compel the state to bring the employment relationship fully into a civil condition. In that way, their disobedience can count as furthering rather than undermining the project of governance by legitimate law. The primary requirements for such principled disobedience are thus that it be oriented toward bringing the employment relationship fully into a civil condition, but without threatening "to usher in a new order of things," and that it target either those who exploit the local illegitimacy (i.e., employers who engage in wage theft) or those who are substantially responsible for the il-

48 Kant, *The Metaphysics of Morals*, 6:257.

49 Kant, *The Metaphysics of Morals*, 6:312.

50 Kant, *The Metaphysics of Morals*, 6:312; see also 6:256.

legitimacy (i.e., those employers, lobbyists, and industry groups working against reform as well as the state itself).

Accordingly, a group of workers might organize a sit-in at their workplace or at a series of workplaces, disrupting normal operations. Or they might go further and organize a longer-term takeover and occupation, stopping normal operations altogether. They may even resist police efforts to remove them, barricading themselves in even if doing so requires damaging or destroying property. They might instead do any of this at the offices of lobbyists hired by employers or industry groups to undermine state- or federal-level efforts at reform, or at the offices of those industry groups themselves. Or they might hack into the computer systems of such lobbying firms or industry groups so as to release to the public internal documents detailing those efforts against reform.

These tactics may have one or more intermediate aims, any of which would be consistent with the larger aim of bringing the employment relationship fully into a civil condition. For instance, their aim may be not only to raise awareness among the wider public of the wage theft they suffer but also to underscore the injustice of it, as the perceived radicalness of their tactics will communicate something about the severity of the injustice. Or their aim may be to raise awareness among other workers, as well as to develop solidarity among themselves, so as to build a wider movement for reform. That the problem of wage theft is not only large but also largely overlooked by the state, the wider public, and even the workers themselves (insofar as they do not recognize it as a systemic problem of nonenforcement) raises the stakes for the workers, making it imperative that they engage in protest actions that will be noticed and taken seriously. Additionally, since employers (and lobbyists and industry groups) are generally politically powerful, the workers' aim in engaging in these kinds of resistance may also be to raise the cost to employers of the status quo, leveraging employers' economic or political self-interest in order, ideally, to force them aside, enlisting them in the cause of reform, or, at the very least, to push them to the sidelines and out of the way of reform.

Workers' resistance may also target the state itself, as it is responsible for their vulnerability to wage theft. Indeed, as we saw, the state's enforcement of employers' rights may even exacerbate this vulnerability. Because of this, it would be consistent with a commitment to the project of governance by legitimate law for workers to engage in disruptive protests and sit-ins of government offices and perhaps even takeovers of those offices tasked with enforcing wage laws. The use of these tactics against the state can have many of the same intermediate aims that their use against employers can have. And here again workers may even resist police efforts to remove them from those offices.

Workers' decisions about whether to pursue these tactics of principled disobedience will be largely strategic decisions. (How likely are these tactics to advance their cause? How likely are they to result instead in counterproductive backlash? Even if they did, might backlash at least build solidarity among workers?) Though these tactics will violate their targets' property rights as defined by the law, these rights do not have the full force of conclusive rights in a civil condition. Workers may use the coercive power available to them to bring about a full civil condition, and these tactics are justified as efforts to do just that.

#### 4. CONCLUSION

Nothing in this discussion is meant to suggest that these tactics are at all likely to be used anytime soon by wage workers in the United States, for there is not a militant labor movement available to organize and pursue these kinds of protest and resistance in a sustained way. But such actions would be justifiable on the Kantian account I have offered, for they would be compatible with a commitment to right. It is thus a mistake to claim, as it seems Kant would, that the only protest available to these workers would be for them to voice their complaints and objections for officials to consider. And it is a mistake because of the kind of defect the American state's failure to enforce their wage rights is: by depriving workers of the equal assurance of the security of their rights that forms part of the basis of the state's claim of authority in the first place, this failure is a defect of legitimacy. And because it is a defect of legitimacy, the Kantian account ought to allow that workers may engage not only in civil disobedience but also in uncivil disobedience, even violent or destructive disobedience, in order to compel the American state to reform itself, for were they to do so, their disobedience would count as furthering rather than undermining the project of governance by legitimate law.<sup>51</sup>

*University of Delaware*  
*dkoltons@udel.edu*

#### REFERENCES

Ashenfelter, Orley, and Robert S. Smith. "Compliance with the Minimum Wage Law." *Journal of Political Economy* 87, no. 2 (April 1979): 333–50.

Bernhardt, Annette, Ruth Milkman, Nik Theodore, Douglas Heckathorn, Mi-

51 I am grateful to two anonymous reviewers for their very helpful comments and criticisms.

- rabai Auer, James DeFilippis, Ana Luz González, Victor Narro, Jason Perlshteyn, Diana Polson, and Michael Spiller. "Broken Laws, Unprotected Workers: Violations of Employment and Labor Laws in America's Cities." National Employment Law Project, 2009. <https://www.nelp.org/publication/broken-laws-unprotected-workers-violations-of-employment-and-labor-laws-in-americas-cities/>.
- Brownlee, Kimberley. *Conscience and Conviction: The Case for Civil Disobedience*. Oxford: Oxford University Press, 2012.
- Colvin, Alexander. "Mandatory Arbitration and Inequality of Justice in Employment." *Berkeley Journal of Employment and Labor Law* 35, nos. 1–2 (2014): 71–90.
- Cooper, David, and Teresa Kroeger. "Employers Steal Billions from Workers' Paychecks Each Year." *Economic Policy Institute Report*, 2017. <https://www.epi.org/125116>.
- Delmas, Candice. *A Duty to Resist: When Disobedience Should Be Uncivil*. Oxford: Oxford University Press, 2018.
- Flikschuh, Katrin. "Reason, Right, and Revolution: Kant and Locke." *Philosophy and Public Affairs* 36, no. 4 (Fall 2008): 375–404.
- . "Sidestepping Morality: Korsgaard on Kant's No-Right to Revolution." *Jahrbuch für Recht und Ethik* 16 (2008): 127–45.
- Galvin, Daniel. "Deterring Wage Theft: Alt-Labor, State Politics, and the Policy Determinants of Minimum Wage Compliance." *Perspectives on Politics* 14, no. 2 (June 2016): 324–50.
- Gilman, Michele. "Supreme Court Ruling against Class Action Lawsuits Is a Blow for Workers—and #MeToo." *The Conversation*, May 21, 2018. <https://theconversation.com/supreme-court-ruling-against-class-action-lawsuits-is-a-blow-for-workers-and-metoo-96976>.
- Gordon, Colin, Matthew Glasson, Jennifer Sherer, and Robin Clark-Bennett. *Wage Theft in Iowa*. The Iowa Policy Project, August 2012. <http://iowapolicyproject.org/2012docs/120827-wagetheft.pdf>.
- Hill, Thomas. "A Kantian Perspective on Political Violence." In *Respect, Pluralism, and Justice*, 33–56. Oxford: Oxford University Press, 2000.
- Holtman, Sarah Williams. "Revolution, Contradiction, and Kantian Citizenship." In *Kant's Metaphysics of Morals: Interpretative Essays*, edited by Mark Timmons, 209–32. Oxford: Oxford University Press, 2002.
- Kant, Immanuel. *The Metaphysics of Morals*. In *Practical Philosophy*, translated and edited by Mary J. Gregor, 363–603. Cambridge: Cambridge University Press, 1996.
- . *Toward Perpetual Peace: A Philosophical Project*. In *Practical Philosophy*,

- translated and edited by Mary J. Gregor, 315–51. Cambridge: Cambridge University Press, 1996.
- Korsgaard, Christine. “Taking the Law into Our Own Hands: Kant on the Right to Revolution.” In *Reclaiming the History of Ethics: Essays for John Rawls*, edited by Andrews Reath, Barbara Herman, and Christine M. Korsgaard, 297–328. Cambridge: Cambridge University Press, 1997.
- Pasternak, Avia. “Political Rioting: A Moral Assessment.” *Philosophy and Public Affairs* 46, no. 4 (Fall 2018): 384–418.
- Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.
- Ripstein, Arthur. *Force and Freedom: Kant’s Legal and Political Philosophy*. Cambridge, MA: Harvard University Press, 2009.
- Shelby, Tommy. *Dark Ghettos: Injustice, Dissent, and Reform*. Cambridge, MA: Harvard University Press, 2016.
- Sinclair, Thomas. “The Power of Public Positions: Official Roles in Kantian Legitimacy.” In *Oxford Studies in Political Philosophy*, vol. 4, edited by David Sobel, Peter Vallentyne, and Steven Wall, 28–52. Oxford: Oxford University Press, 2018.
- Stilz, Anna. *Liberal Loyalty: Freedom, Obligation, and the State*. Princeton: Princeton University Press, 2009.
- . “Provisional Right and Non-State Peoples.” In *Kant and Colonialism: Historical and Critical Perspectives*, edited by Katrin Flikschuh and Lea Ypi, 197–220. Oxford: Oxford University Press, 2014.
- US Department of Labor. *The Social and Economic Effects of Wage Violations: Estimates for California and New York*. Prepared by Eastern Research Group, 2014. <https://www.dol.gov/asp/evaluation/completed-studies/WageViolationsReportDecember2014.pdf>.



## ON EMAD ATIQ'S INCLUSIVE ANTI-POSITIVISM

Kara Woodbury-Smith

THERE ARE well-known instances of morally abhorrent law, like the legal rules of Nazi Germany or apartheid South Africa. According to Emad Atiq, the existence of morally abhorrent legal rules presents an *extensional challenge* to legal anti-positivism.<sup>1</sup> Atiq acknowledges that morally abhorrent legal rules are intuitively legally valid, and yet anti-positivism (which maintains that the legal validity of a rule is partly grounded in that rule's moral merit) cannot explain why. In his paper, "There Are No Easy Counterexamples to Legal Anti-positivism," Atiq puts forward a type of anti-positivism capable of responding to this specific extensional challenge. He calls it Inclusive Anti-positivism (IAP).

IAP "identifies the property of *being law* with the ... normative property of rules (*being normatively well supported to a high enough degree*)."<sup>2</sup> Atiq's reasoning goes like this: we can identify legal rules wherever the subjects of those rules have "broadly normative reasons" to follow them.<sup>3</sup> These broadly normative reasons are grounded in irreducible moral facts and, as such, the legal validity of rules is partly grounded in the existence of moral facts that are not dependent on social facts. Because IAP grounds legal validity in a less restrictive conception of moral facts (we are no longer concerned with the moral merit of the rule in question when determining its legal validity, but whether it is essentially good to follow), IAP is able to account for the legality of morally abhorrent rules while preserving morality as an existence condition of law.

In section I, I point out that IAP presupposes a conceptually necessary connection between law and coercion. In section II, I briefly discuss internal- and external-to-practice appraisals of legal rules. Last, in section III, I touch upon the explanations of legal normativity offered by IAP and some positivists. The point of this response is not to deny Atiq's claim that there can be moral reasons to comply with morally abhorrent law. Rather, its aim is to raise a question: Is IAP's

1 Atiq, "There Are No Easy Counterexamples to Legal Anti-positivism," 3 (hereafter cited as "Counterexamples").

2 Atiq, "Counterexamples," 16.

3 Atiq, "Counterexamples," 13.

central claim about legal validity necessarily at odds with positivism's central claims that law is fundamentally a social institution and that what counts as law ultimately depends on social sources?<sup>4</sup>

## I

Rules are normative. This means that they are action guiding—they serve as reasons for action and as a means to appraise the conduct of others. Take netball: if you want to play netball, then you have reason to comply with the rules of netball. If you do not follow the rules, then others can say that you are playing netball incorrectly and not pick you for their team. The law is taken to be normative because, within a given territory, legal rules typically tell law subjects what they must refrain from doing and how to do things legally (e.g., how to enter into a contract).

Unlike the rules of netball, law's normativity is taken to be unconditional. Law does not say: if you want to be law abiding, then you should not *A*. Rather, it says to everyone within its jurisdiction, regardless of whether they accept the law as an authority, "*Do not A*." This "conditionality" of rules leads to a distinction: that there are two basic senses of normativity.<sup>5</sup> That my daughter, who cannot swim, will drown if she goes into a pool alone is an *objective* reason for her to comply with my rule, "*Do not go swimming alone*." This means that her inability to swim is a reason to "*not swim alone*" that exists even if she thinks that the water is shallow or that she is secretly a mermaid. On the other hand, her belief that she is secretly a mermaid is also a reason for action, but in the *subjective* sense. Despite what objectively may be the case (she cannot swim), she thinks she has reason to go swimming alone (she is secretly a mermaid).

Atiq's General Grounding Claim (GGC) for legal validity asserts: "the rule's legality is grounded in whatever normative reasons there are for agents to follow the rule."<sup>6</sup> So, per IAP, a rule can only be legally valid if there is an *objective* reason for complying with it. According to Atiq, we have objective reasons for complying with a rule whenever the rule in question surpasses a normative threshold that makes it worthy of adoption into the legal system by relevant officials.<sup>7</sup> The normative threshold that marks the difference between legally valid and not legally valid can be surpassed in three situations: (1) if a rule is of a certain moral

4 Sometimes this claim is known as the "Social Thesis" or "Social Fact Thesis." See Himma, "Philosophy of Law"; and Woodbury-Smith, "Inclusive Legal Positivism."

5 For a classic discussion of the two senses in which rules apply, see Foot, "Morality as a System of Hypothetical Imperatives."

6 Atiq, "Counterexamples," 16.

7 Atiq, "Counterexamples," 20–22.

standard; (2) if a rule is an entrenched social convention; or (3) if a rule is a blend of the two.<sup>8</sup> This means there can be rules whose legal validity is rooted in their moral merit, even if they are not conventionally followed. Likewise, there can be rules whose legal validity is grounded in their existence as entrenched social conventions, even if they are morally abhorrent.

IAP, therefore, shifts the focus of anti-positivism away from considerations of a rule's moral merit (the focus of classic anti-positivism) and toward the objective reasons we, as subjects, have for complying with rules, including the morally abhorrent ones. Critically for Atiq, such reasons are *always* moral. Much of his article is devoted to supporting the assertion that weakly moral reasons for action exist, and that self-protection is a weakly moral reason for action.<sup>9</sup> So, if there are weakly moral reasons like self-protection to comply with a morally abhorrent law, then that law can satisfy the GGC for legal validity and IAP is counterexample proof.

Let us grant Atiq's GGC, that what makes a morally abhorrent rule legally valid is that it is socially entrenched to such a degree that its subjects have objective moral reason(s) to comply with it. Is it the case that this conception of legal validity is actually *anti*-positivist in that it is not ultimately grounded in social facts? The answer here must be no.

Consider Nazi Germany. According to Atiq, while there were overwhelmingly strong moral reasons for its subjects to reject and subvert Nazi Germany's laws, there was also at least one weakly moral reason to comply with them: the good of self-protection.<sup>10</sup> Because of this weakly moral reason to comply with Nazi law, Nazi law was "good to follow to some degree."<sup>11</sup>

What does it take for Nazi law to be "good to follow to some degree"? According to Atiq, it is when "deviating from conventionally embraced rules renders individuals vulnerable to sanction."<sup>12</sup> He later writes that the "moral reason to follow a conventionally embraced rule might be partly grounded in nonmoral facts, like the rule's conventionality. But it is also grounded in a pure moral fact: the moral principle that *if following a rule promotes your interests*, then there is some moral reason to follow it."<sup>13</sup>

Here I want to point out that there can be a difference between complying

8 Atiq, "Counterexamples," 15–16. Throughout the discussion, when I refer to morality I am invoking a sense of ideal morality.

9 Atiq, "Counterexamples," 6–8.

10 Atiq, "Counterexamples," 6–8.

11 Atiq, "Counterexamples," 9.

12 Atiq, "Counterexamples," 6.

13 Atiq, "Counterexamples," 13, emphasis added.

with a law for self-protection (e.g., behaving legally because otherwise one would be liable to sanction) and complying with a law for the promotion of one's interests. Consider a law (valid in accordance with its sources) requiring that I throw rocks at unwed mothers. Further consider that this law is conventionally practiced. In the absence of institutionalized sanctions, not complying with this law would, at worst, decrease foot traffic to my business and suppress my interest in having lots of money. Now others who prioritize having lots of money above their neighbor's health may disagree, but it is hard to see how I could have even a weakly moral reason to comply with that law when the worst I would face for refusing to cast stones upon certain persons is a reduction in my wealth.

However, if the legal system backed this law with serious sanctions and my noncompliance meant that the state would seize my business and detain my husband, my children, and myself in labor camps, *then* it seems like I could have a weakly moral reason to comply with the law, as such compliance is about promoting my interests in survival and protecting my family. Indeed, this was how Nazis were able to create Jewish police collaborators in their ghettos: by appealing to our innate drive to protect ourselves and our families.

Atiq's metaethical theorizing may check out. Self-protection may be an objective, weakly moral reason for complying with morally abhorrent rules. However, it seems to lead IAP into a conundrum. Following IAP, one can only appeal to self-protection as a moral reason for complying with morally abhorrent rules if certain nonmoral social facts obtain, like institutionalized sanctions. If that is correct, then the legal validity of morally abhorrent rules, per IAP, ultimately hangs on contingent social facts—specifically, whether the legal system backs such rules with serious sanctions. Atiq has, it seems, developed an anti-positivist command theory of law. Atiq may wonder how significant this observation is, given there are not, at least to my knowledge, any instances of morally abhorrent law that are not backed by sanctions. However, if IAP is meant to include a concept of legal validity in general (like positivism does), then it is an observation that needs consideration.

## II

Atiq writes: "If inclusive anti-positivism . . . is true, it remains possible for a judge to comply with her legal duties while striking down morally abhorrent laws for conflict with other laws that are morally optimal even if weakly conventional."<sup>14</sup> As an example, Atiq discusses the moral rule requiring respect for human dignity. Such a rule has "enough morally going for it" that it meets the normative

<sup>14</sup> Atiq, "Counterexamples," 22.

threshold for legal validity in Nazi Germany.<sup>15</sup> Atiq goes on to write: “Had [a rule requiring respect for human dignity] been recognized as law by Nazi jurists, there might have been greater official resistance against the Third Reich.”<sup>16</sup>

The idea here is that any of the Nazi jurists could have criticized Nazi law for being inconsistent with the rule requiring respect for human dignity for all (which, per IAP, is legally valid because it is of a certain moral standard) and still be complying with their legal duty to interpret and apply the law.<sup>17</sup> However, none of them did and the IAP claim that there was a legally valid rule requiring respect for human dignity for all in Nazi Germany strikes me as being at odds with the same intuitions Atiq relies upon when he appeals to the legal validity of morally abhorrent rules. Atiq writes that denying that the Nazis had law is counterintuitive.<sup>18</sup> How is it not similarly counterintuitive to claim that Nazis had a legal rule requiring respect for human dignity for all? The moral rule requiring respect for human dignity was not a legally valid rule in Nazi Germany for the simple reason that it was never recognized as such by any relevant legal official, despite its moral worth.

Incorporating the feature *being morally good to follow* as a conceptually necessary feature of law is sold as a benefit because it allows for an “internal-to-practice moral critique of a legal system” such that all moral critique of law is legal critique.<sup>19</sup> IAP can better track what Atiq calls the “judicial intuitions about the legality of rules” because it supports the following legal reasoning: “posited law  $x$  was never actually legally valid (or, is now invalidated) because it contravenes non-posited law  $y$ ” (where  $y$  meets the normative threshold for legal validity under IAP on the basis of its moral worth).<sup>20</sup> Inclusive positivists, Atiq acknowledges, have the ability to draw on constitutional standards and engage internal-to-practice moral critiques of law. The problem, for Atiq, is that this ability is contingent on social facts.<sup>21</sup> However, this observation is not entirely correct, as positivists, inclusive and exclusive, can engage in internal-to-practice moral appraisal of law because drawing on rule-of-law principles is always a live option.

I wonder what, exactly, the benefit is of a conception of law that concludes

15 Atiq, “Counterexamples,” 21.

16 Atiq, “Counterexamples,” 22.

17 Atiq, “Counterexamples,” 16.

18 Atiq, “Counterexamples,” 3.

19 Atiq, “Counterexamples,” 22.

20 Atiq, “Counterexamples,” 18.

21 Atiq, “Counterexamples,” 22. If this is problematic, then it should also be problematic for Atiq that IAP’s ability to account for the legal validity of morally abhorrent rules is similarly contingent on social facts.

that all moral critique of law is legal critique?<sup>22</sup> The “familiar rhetoric of the judicial process” may be such that when some judges morally critique the law, they take it to be a legal critique.<sup>23</sup> However, I see no practical difference between the following decisions: (1) “Legal rule  $x$  should be struck down because it contravenes moral rule  $y$ ” or (2) “Legal rule  $x$  should be struck down because it contravenes legal rule  $y$ ” (where  $y$ ’s legality is entirely a function of its being moral principle  $y$ ). In either case, if the relevant officials buy the arguments about  $x$ , and their legal authority grants them the relevant powers, then  $x$  will be struck down. Decision 2 may allow judges to say to themselves that they engaged solely in a legal critique of the law, but that is smoke and mirrors.

One of the most interesting parts in *The Concept of Law* comes in the penultimate chapter, where H. L. A. Hart writes:

What surely is most needed in order to make men clear sighted in confronting the official abuse of power, is that they should preserve the sense that the certification of something as legally valid is not conclusive of the question of obedience, and that, however great the aura of majesty or authority which the official system may have, its demands must in the end be submitted to a moral scrutiny.<sup>24</sup>

Hart is speaking to the power that can come from the separability of moral merit from legal validity. Our laws are *always* capable of being morally scrutinized, externally to legal practice, not only by our officials, but also by us—and this is a good thing.

### III

Another area where Atiq sees a key explanatory difference between IAP and positivism is in their explanations of law’s normativity. Atiq writes that “no positivist, as far as I can tell, construes the property of legality as essentially identical to a *bona fide* normative property.”<sup>25</sup> He is correct: positivists do not typically make strong assertions as to whether a legal rule is objectively normative in virtue of its legal validity and this is simply because positivists do not have to.

22 Atiq claims that “one of the principal motivations for being an anti-positivist is the possibility of a moral critique and improvement of law from an *internal-to-law* perspective” (“Counterexamples,” 21).

23 Hart, *The Concept of Law*, 274.

24 Hart, *The Concept of Law*, 210.

25 Atiq, “Counterexamples,” 17. For an example of a positivist doing just this, see Himma, “A Comprehensive Hartian Theory of Legal Obligation” and *Coercion and the Nature of Law*.

Here is one way a positivist could explain the normative property of a legal rule: the traditional Razian assertion is that legal rules *claim* to provide us with objective reasons for action.<sup>26</sup> This means that there can be instances in which legal rules fail to provide us with such reasons. Put far too simply: in the Razian scheme, legal rules claim to be first- and second-order exclusionary reasons for action.<sup>27</sup> Imagine that you want to *y* because *y* is fun. Act *y* is morally inert and is legal where you live, but you are currently on vacation in another country, where, you discover, there is a law prohibiting *y*. That legal prohibition claims to give you a first-order reason not to *y*. This means it claims to be a reason in itself: you should not *y* because the law you are currently subject to prohibits it. The legal prohibition also claims to be a second-order exclusionary reason for action. This means that, because you are subject to a law prohibiting *y*, you should not act in accordance with other reasons you have to *y*, like the prudential reason “*y* is fun.”

There is a sense of the term “obligation” according to which to be under a legal obligation simply means that one is subject to a mandatory rule—a rule that *claims* to impose an obligation. This is the sense we can have in mind with regard to the legal rule in apartheid South Africa that prohibited Black and white persons from associating with one another on beaches. Because of that legal rule, there was a legal obligation not to associate. And yet the moral repugnance of that legal rule and the ideology it sustained entailed that it actually provided no objective reason for action, save the avoidance of sanctions. So, despite what some legal rules may *claim* (that they are first- and second-order reasons for action), it is possible for them to fail to *actually* provide such reasons for action.

What does this have to do with IAP? As stated, IAP asserts that legal validity of a rule is identical to the rule's property of being actually normative—of actually providing objective reasons for action. Where morally abhorrent legal rules are concerned, their normative property comes down to socially contingent considerations raised in section 1: whether such rules are backed by sanctions. So, per IAP, in apartheid South Africa the desire to avoid sanctions gave white and Black persons objective reason to comply with its morally abhorrent legal rules. What I hope I have made clear in this section is that this is the same explanation of legal normativity offered by the Razian sketch above: even if a legal rule does not itself give its subjects objective reasons to comply with its directives, the legal system may give other reasons for compliance by, for instance, threatening the use of sanctions.

The only relevant difference between IAP and the Razian picture is the connection between legal validity and the normative property of legal rules. For

26 Raz, *The Authority of Law*, 30.

27 Raz, *Practical Reason and Norms*, esp. 35–48, 58–59.

those who accept the Razian picture, inquiries into legal validity are distinct from inquiries into the reasons we have for complying with legal rules. Atiq may be generally correct when he notes that positivists do not typically construe legal validity as identical to its objective normative property, but that is simply because positivists can explain the legal validity of a rule without needing to draw such an identity claim. To connect this point back to the passage by Hart at the end of section II, there are very good reasons for preferring such a conception of law.

## IV

With IAP, Atiq has shifted the moral fact of interest away from the moral merit of a legal rule and toward the rule's property of *being good to follow*. This is how IAP is able to account for the legal validity of morally abhorrent rules. In this discussion I raise three observations. First, our ability to appeal to weakly moral reasons to comply with morally abhorrent law is contingent on social facts (e.g., that the law is backed by sanctions).<sup>28</sup> Second, the upshot to IAP that all moral appraisal of legal rules is legal appraisal is not necessarily as appealing as Atiq takes it to be. Last, the Razian thesis that law *claims* to provide its subjects with objective reasons for action is not necessarily at odds with the picture of legal normativity that Atiq draws.

Of the three observations, I think the first presents the most trouble for IAP. Positivists argue that law is a social institution and that what counts as valid law ultimately depends on social sources. As I observed in section I, following IAP, unless the state gives its subjects weakly moral reasons to comply with its rules (by institutionalizing sanctions), then those rules, should any of them be morally abhorrent, would fail to be legally valid. Given this essential connection between the contingent social fact of sanctions and legal validity, IAP seems to be grounded by the positivist principle regarding the social nature of law. And so, is IAP actually *anti-positivist*?<sup>29</sup>

Durham University  
kara.m.woodbury-smith@durham.ac.uk

28 Atiq could perhaps make the claim that obedience to morally abhorrent law is morally required for the common good, which is Aquinas's view (though Aquinas, of course, maintains that morally abhorrent human laws are not genuine instances of law). See Aquinas, *Summa Theologica*, q.96, a.4.

29 I am indebted to Emad Atiq who wrote an article that I thoroughly enjoyed spending time with (no small thing as this was written during the first lockdown of the COVID-19 pandemic). My thanks as well to the anonymous reviewers for their thoughtful and supportive com-



## REFERENCES

- Atiq, Emad. "There Are No Easy Counterexamples to Legal Anti-positivism." *Journal of Ethics and Social Philosophy* 17, no. 1 (March 2020): 1–26.
- Aquinas, Thomas. *Summa Theologica*. Translated by Fathers of the English Dominican Province. New York: Benziger Brothers, 1911–1925.
- Foot, Philippa. "Morality as a System of Hypothetical Imperatives." In *Virtues and Vices*, 305–16. Berkeley: University of California Press, 1978.
- Green, Leslie. "Positivism and the Inseparability of Law and Morals." *New York University Law Review* 83, no. 4 (October 2008): 1035–58.
- Hart, H. L. A. *The Concept of Law*. 3rd ed. Oxford: Oxford University Press, 2012.
- Himma, Kenneth Einar. *Coercion and the Nature of Law*. Oxford: Oxford University Press, 2020.
- . "A Comprehensive Hartian Theory of Legal Obligation: Social Pressure, Coercive Enforcement, and the Legal Obligation of Citizens." In *Philosophical Foundations of the Nature of Law*, edited by Wil Waluchow and Stefan Sciaraffa, 152–82. Oxford: Oxford University Press, 2013.
- . "Philosophy of Law." *Internet Encyclopedia of Philosophy*. <https://iep.utm.edu/law-phil/>.
- Raz, Joseph. *The Authority of Law: Essays on Law and Morality*. 2nd ed. Oxford: Oxford University Press, 2009.
- . *Practical Reason and Norms*. London: Hutchinson and Co., 1975.
- Woodbury-Smith, Kara. "Inclusive Legal Positivism." In *Encyclopedia of the Philosophy of Law and Social Philosophy*, edited by Mortimer Sellers and Stephan KIRSTE. [https://doi.org/10.1007/978-94-007-6730-0\\_62-2](https://doi.org/10.1007/978-94-007-6730-0_62-2).

---

ments on an earlier draft, and to Margaret O'Brien, Johanna Jacques, and Wil Waluchow for their encouragement and insights.