

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XXIII · NUMBER 1

October 2022

ARTICLES

Don't Be Cruel: Building the Case for Luck in
the Law

Alexander Sarch

Expecting Equality: How Prenatal Screening
Policy Harms People with Disabilities

Athmeya Jayaram

Children, Partiality, and Equality

David O'Brien

Democracy and Social Equality

Ryan Cox

What Goes On When We Apologize?

Christopher Bennett

DISCUSSIONS

Against Being For

James L. D. Brown

No Grit without Freedom

Berislav Marušić

JOURNAL *of* ETHICS
& SOCIAL PHILOSOPHY

VOLUME XXIII · NUMBER 1

October 2022

ARTICLES

- 1 Don't Be Cruel: Building the Case for Luck in
the Law
Alexander Sarch
- 32 Expecting Equality: How Prenatal Screening
Policy Harms People with Disabilities
Athmeya Jayaram
- 59 Children, Partiality, and Equality
David O'Brien
- 86 Democracy and Social Equality
Ryan Cox
- 115 What Goes On When We Apologize?
Christopher Bennett

DISCUSSIONS

- 136 Against Being For
James L.D. Brown
- 144 No Grit without Freedom
Berislav Marušić

JOURNAL of ETHICS & SOCIAL PHILOSOPHY
<http://www.jesp.org>

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.

Executive Editor

Mark Schroeder

Associate Editors

Saba Bazargan-Forward	Hallie Liberto
Stephanie Collins	Errol Lord
Dale Dorsey	Tristram McPherson
James Dreier	Colleen Murphy
Julia Driver	Hille Paakkunainen
Anca Gheaus	David Plunkett

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

Rachel Keith

Copyeditor

Susan Wampler

Typesetting

Matthew Silverstein

DON'T BE CRUEL

BUILDING THE CASE FOR LUCK IN THE LAW

Alexander Sarch

LUCK ABOUND in the law. Suppose Jack and Jill both drive 55 mph in a 25 mph zone while aware that doing so imposes a serious risk of harm. Jack gets lucky and does not hit anyone, while Jill hits and kills a ten-year-old who suddenly darts into the road. Many jurisdictions have no general crime of reckless endangerment, which would make one criminally liable even if no harm eventuates provided one behaved in ways one was aware imposed a substantial and unjustified risk of harm.¹ In jurisdictions lacking a general crime of reckless endangerment (and no serious inchoate traffic crimes), Jack would not face severe criminal liability; he would at most have committed routine traffic violations. Jill, by contrast, would face far more serious criminal sanctions.

What has proven persistently puzzling about this feature of the law is that Jill would face greater criminal liability than Jack *even if* both are equally culpable for their conduct—that is, their actions both manifest the same amount of insufficient regard for the legally protected interests of others.² They plausibly are equally culpable if we suppose they acted the same way with the same beliefs and intentions and ended up causing different levels of harm only due to unexpected events over which they had no control. Assuming the amount of criminal liability one deserves tracks the culpability of one's conduct, Jack

1 As Kaplan, Weisberg, and Binder note, “we rarely punish harmless negligence or recklessness (charges of ‘reckless endangerment’ are uncommon . . .), and where we do . . ., we usually . . . narrowly define the risky conduct that will give rise to liability (as in drunk-driving laws), rather than just proscribing any conduct that poses an unacceptable risk to a particular interest” (*Criminal Law*, ch. 10). Even fewer jurisdictions recognize negligent conduct as a general crime where no harm eventuates. Thus, a similar puzzle arises for negligent conduct as well.

2 As Alexander and Ferzan put it, “insufficient concern [is] the essence of culpability” (*Crime and Culpability*, 67–68). For other defenders of the insufficient regard theory of culpability, see Tadros, *Criminal Responsibility*, 250; Westen, “An Attitudinal Theory of Excuse,” 373–74; and Sarch, *Criminally Ignorant*, 27–64.

and Jill deserve to be punished the same. But Jill may still be punished more harshly. How can this be justified?

This question of how to justify the criminal law's recognition of luck—the *legal luck puzzle*—is not limited to endangerment. It also arises for attempt, which is a crime of intent. Suppose Bert and Ernie both fire their guns at a person with the intent to kill, but one of Bert's pet pigeons unexpectedly swoops down and blocks Bert's bullet. Bert's attempt fails, while Ernie's succeeds. In many jurisdictions, Ernie will face greater criminal liability than Bert.³ In these jurisdictions, a more serious label will be applied to Ernie—"murderer"—and he is liable to be punished more harshly. Nonetheless, Bert plausibly is just as culpable for his conduct as Ernie is for his: they both intended the death and believed it was practically certain to occur. Here, I simply assume such pairs of actors (like Bert and Ernie or Jack and Jill) are identical in all respects—including culpability—except that one happens to cause harm while the other does not.⁴ Thus, the legal luck puzzle arises here too. How could it make sense to impose greater criminal liability on Ernie than Bert, even though they are assumed to be equally culpable for what they did? Is this not paradigmatic injustice—the unequal treatment of actors who differ in no normatively relevant respect?

Many solutions to the legal luck puzzle have been proposed. David Lewis sought to explain the differential treatment of such actors as a justifiable form of

- 3 LaFave's criminal law treatise observes that many jurisdictions provide lesser punishment for attempts than for the analogous completed crimes, although the manner in which this is done varies. As he notes, many "modern recodifications . . . [declare] the attempt to be a crime one degree below the object crime," and for "statutes dealing with attempts to commit particular crimes, the authorized punishment is usually lower than for the completed crime" (*Substantive Criminal Law*, sec. 11.5). See also Christopher, who notes that "almost every jurisdiction world-wide punishes the attempt that succeeds more severely than the attempt that fails" ("Does Attempted Murder Deserve Greater Punishment than Murder?" 419). However, LaFave also notes that this practice is not universal: some provisions "provide that the penalty for attempt may be as great as for the completed crime" (*Substantive Criminal Law*, sec. 11.5). Cf. American Law Institute, *Model Penal Code*, sec. 5.05(1), which subjects attempts to the same penalties as the completed crime, except for attempts to commit a capital crime or a felony of the first degree, which is an offense of the second degree. The present paper is concerned with why jurisdictions that *do* impose lower penalties on attempts might be justified in doing so, even if other jurisdictions might reasonably choose a different course.
- 4 This is a common position to take on such cases as this. See Lewis, "The Punishment That Leaves Something to Chance"; Edwards and Simester, "Crime, Blameworthiness, and Outcomes"; and Sarch, *Criminally Ignorant*, 81n128. Some might deny the existence of the puzzle by claiming that causing harm makes one more culpable. However, in what follows, I will set aside this response and proceed on the assumption that in a pair of cases like the above the two actors are equally culpable.

penal lottery.⁵ James Edwards and Andrew Simester argue that the person who unluckily causes harm *did something different*, and thus may be guilty of a more serious offense, even if she is no more culpable than the analogous attempter who avoids harm through mere luck.⁶

The aim of this paper is not to evaluate these responses to the legal luck puzzle. Instead, I develop a new sort of nonconsequentialist argument that justifies the criminal law's imposition of differential treatment on equally culpable actors who differ only as to the harm caused. Importantly, however, my argument does not *mandate* a policy of differential punishment for harmful and harmless wrongdoers. My aim is more modest: to provide a rationale that a legislature in jurisdictions like ours can use to normatively justify its decision to adopt such a policy. I do not claim that the decision to recognize luck is the *only* way to comply with the relevant normative principles. Instead, all I argue is that there is robust normative support for this decision compared to other natural ways to satisfy the underlying normative principles.

The sort of argument I focus on does not ask what reason there is to *ratchet up* the criminal liability of harmful actors (Ernie, Jill) above the level faced by the analogous luckily harmless actors (Bert, Jack). Instead, it asks what reason there is to *ratchet down* the criminal liability of luckily harmless actors. The sort of justification I am interested in starts by assuming the criminal law would be permitted to impose just as much liability on Bert (the attempter) as it does on Ernie (the murderer), and to treat Jack (the lucky reckless driver) the same as Jill (his lethal counterpart) because both pairs of actors are assumed to be equally culpable for their conduct. Still, ratcheting-down arguments provide reasons for the legislature to treat harmless actors less harshly than it is entitled to.⁷

Thus, instead of focusing directly on features of the wrongdoer, ratcheting-down arguments focus on reasons the *legislature*, when passing criminal laws, should take into account. Shifting focus to the legislature's standpoint, I suggest, is a promising avenue for securing a justification of luck in the law.

In section 1, I critique existing ratcheting-down arguments. They tend to rely on *pragmatic* reasons for lowering the criminal liability of less harmful

5 Lewis, "The Punishment That Leaves Something to Chance."

6 Edwards and Simester, "Crime, Blameworthiness, and Outcomes."

7 Similar arguments can also explain why someone who causes a smaller amount of harm (say, \$500 of property damage) merely by luck might justifiably be punished less than analogous equally culpable actors who cause more harm (say, \$5,000 of damage). However, it should be obvious how my arguments carry over to these cases. For clarity, therefore, I primarily focus on why it makes sense to punish *harmless* actors less harshly than equally culpable harmful actors.

actors. Ideally, however, we want a genuinely *normative* ratcheting-down argument—one that is not similarly held hostage to practical limitations or unstable factual circumstances. I focus instead on arguments that are more nonconsequentialist in nature.⁸

Thus, in section 2, I develop a new ratcheting-down argument—a more principled justification for countenancing some luck—while section 3 considers alternative solutions. My argument starts from the idea that the legislature, when passing criminal laws, has a duty not to be heartless, mean, vicious, and vindictive—what I dub *the duty to not be cruel*.⁹ I argue that the legislature would breach this duty if it always passed laws that criminalize conduct to the full extent permitted on culpability grounds. Withholding some punishment that would otherwise be due to harmless wrongdoers in virtue of their culpability is a particularly apt way for the legislature to comply with this duty. The upshot is a promising justification for imposing less criminal liability for culpable but luckily harmless conduct. The argument does not aim to show that recognizing luck is *required*. A legislature may adopt other institutional arrangements to avoid the relevant form of cruelty. Still, I argue this duty offers an adequate normative basis for the legislature to withhold some punishment otherwise due to wrongdoers who cause less harm. In closing, I address worries about whether my argument leads to a conflict with retributive justice.

One caveat. While the argument I develop could in principle be used to justify any way that luck impacts criminal liability, it does not have to. Some might think that the normative claims underlying the argument are stronger for some crimes than others. In what follows, I explain how the argument might apply in different ways to endangerments and attempts. My primary concern will be to defend the argument as to endangerment, though I also show how the argument can be modified to extend to attempts.

8 My argument also rests on some contingent facts—e.g., assumptions about human psychology. See *infra* text accompanying notes 32–33. If human psychology were very different, the argument might not hold. Nonetheless, the factual circumstances I rely on are more fundamental and stable background conditions compared to the highly changeable facts appealed to in the arguments I criticize in section 1. This is only a difference in degree, not in kind. But the extra stability and broader applicability of the assumptions behind my argument should make it more interesting. (Thanks to Joe Horton and Erik Encarnacion for this point.)

9 Rawls recognizes a duty not to be cruel and vindictive, which he deems a “natural” duty (i.e., applicable apart from its being endorsed by those in the original position) (*A Theory of Justice*, 114).

1. UNSATISFACTORY RATCHETING-DOWN ARGUMENTS

By a “ratcheting-down argument,” I mean one that explains why the criminal law is justified in lowering the amount of criminal liability imposed on less harmful actors like Bert and Jack, even though they are just as culpable as their more harmful counterparts, Ernie and Jill. What is the best way to construct such a ratcheting-down argument?

1.1. *Self-Interested Ratcheting-Down Arguments*

One version is pragmatic. The government might have self-interested reasons not to punish as harshly when there is no easily identifiable victim who was harmed.¹⁰ Some might think harmful crimes seem easier for law enforcement to detect and prosecute than nonharmful conduct. When harm results, it is more likely there will be a victim or other parties who are in a position to alert the authorities. Accordingly, the legislature might decide that law enforcement will be more cost effective if it focuses on misconduct that tends to be easier to detect.¹¹

This argument is not convincing. Most importantly, nonharmful conduct—whether attempt or endangerment—often *is* easy enough to detect. There may be multiple witnesses to Jack’s driving 55 mph in a 25 mph zone who alert the police. So why not criminalize all inchoate recklessness and allow law enforcement to decide how best to allocate its investigative resources?

Consider, therefore, a second pragmatic argument, based on political self-interest.¹² The legislature might reason that it is likely to remain more popular if it punishes only to the extent it “has to”—that is, only when *not* punishing more harshly would be politically costly. Perhaps being as aggressive as possible in its criminalization choices might alienate voters whose friends and family are imprisoned. To remain popular, the legislature might prefer to impose less harsh penalties. Of course, at some point, *failing* to punish more harshly would also entail political costs. Thus, to strike a balance, the legislature might decide to punish only to the extent needed to avoid political repercussions. So even when actors like Bert and Jack are culpable, the legislature might

10 Cf. Lewis, “The Punishment That Leaves Something to Chance,” 54, which discusses the idea that when the public does not “see blood” they perhaps will not “demand blood.”

11 Cf. Simester, “Is Strict Liability Always Wrong?” 47, which notes that “this requirement [that harm should occur before risky conduct is criminalized] would also have the practical advantages of reducing the costs of policing.”

12 See Sarch, “Knowledge, Recklessness and the Connection Requirement between Mens Rea and Actus Reus,” 34n97, which discusses this idea in another context.

prefer not to punish (or punish less) because it does not *have to* in order to avoid political costs.

Still, there are problems. The first is dependency on highly contingent facts. The present rationale applies only where the right political attitudes prevail. In many jurisdictions, political costs are more likely for underpunishment, as it invites criticism for being “soft on crime.” It is more likely the legislature’s self-interested calculus would support imposing *more* criminal liability to the extent it can get away with it.

Moreover, arguments from self-interest are the wrong kind of reason.¹³ They do not provide a *normative* justification for the criminal law to countenance luck. Self-interested reasons for the legislature to withhold some punishment do not show why this would be *normatively justified*—that is, fair, just, or otherwise morally defensible. Related concerns afflict the first pragmatic argument sketched, based on making law enforcement more cost effective. Considerations of cost legitimately bear on the justification of legislative policies, but I am seeking nonconsequentialist, fairness reasons for the content of the criminal law to recognize luck.

1.2. Mercy

A better ratcheting-down argument stems from the idea that the legislature can display the virtue of mercy by declining to impose as harsh punishments on those who fortuitously cause less harm as on the analogous, equally culpable actors who cause more harm. Since it is good to display mercy in general, the legislature would have a reason to do so when passing criminal laws in particular. This gives another argument for imposing less criminal liability on Bert and Jack than Ernie and Jill, respectively.

This argument has advantages over the previous one. It does not depend on highly contingent facts about what preserves government popularity. Rather, it is a reason for ratcheting down that always applies—regardless of political climate, limited resources, or similarly changeable circumstances. Moreover, it more plausibly is the right kind of reason.

Nonetheless, this argument has problems. First, it is unclear why mercy should routinely be shown to nonharmful actors but not to harmful ones. Insofar as there is a general reason to show mercy, would it not apply in all cases regardless of harm? True, the legislature might decide that if mercy is to be shown at all, it is less likely to arouse controversy if shown to the harmless. After all, the population is more likely to demand a strong state response when harm

13 Cf. Lewis, “The Punishment that Leaves Something to Chance,” 54, which observes that the rationale that the public demands blood only when it sees blood “does nothing at all to defend our practice as just.”

is caused. Still, answering this way is to import pragmatic considerations of the kind that we just saw to be reasons of the wrong kind.

There is also a deeper problem. Mercy is *optional*.¹⁴ It is not generally something one *must* do. It might be desirable for the state to show mercy—just as it might be desirable for the state to display other virtues like courage or good taste. We might want the state to manifest such traits. But declining to show mercy, I take it, is not grounds for complaint by the person affected. Mercy is not something the legislature must do on pain of being an apt target of complaint or blame. It is not something the legislature has a *duty* to display.

As a result, the mercy argument does not appear weighty enough to account for the wide range of cases in which legal systems punish harmless conduct less harshly—that is, both for attempts and for harmless reckless conduct. Consider all the scenarios where legal systems like ours refrain from punishing less harmful actors as much as more harmful, equally culpable actors. Can all of these really be explained as a form of merely optional mercy? Perhaps *some* might. But since mercy is optional, we would expect the reasons to display mercy when legislating in criminal contexts to often be overridden by other reasons the legislature must consider—like deterrence and the value of expressing condemnation of culpable conduct. It is doubtful that mercy is a sufficiently weighty consideration to account for this broad range of cases. Accordingly, we should keep looking for reasons that more fully justify the legislative decision to punish less harmful actors less harshly—reasons that are not optional the way mercy is, but rather are more firmly grounded in the duties of the legislature. (I revisit the comparison to mercy in section 3.2.)

1.3. Reducing the Risk of Abuses

Another ratcheting-down argument is that taking harm to be a prerequisite for punishment might help reduce the risk of abuses by law enforcement. Andrew Simester has suggested a similar idea in another context.¹⁵ Here, the thought is that a simple way to reduce the risk of police and prosecutors abusing their discretion would be to restrict the amount of conduct that is criminalized. Having fewer or narrower crimes on the books would restrict the scope of behavior that

14 See, e.g., Tasioulas, “Mercy,” which expresses skepticism (in section 4) about the idea that there is a duty to show mercy or that offenders have a right thereto.

15 In discussing justifications for strict liability as to result elements, Simester suggests that declining to criminalize risky conduct unless harm ensues “has rule of law benefits, since it confines what otherwise might be wide-ranging discretionary powers of arrest and prosecution. . . . One way of minimizing the risk of malicious or discriminatory prosecutions would be to qualify the offence by [including] a strict liability requirement” of harm to a victim (“Is Strict Liability Always Wrong?” 47).

police and prosecutors are authorized to intrude into. Declining to criminalize harmless conduct would help restrict the domain of behavior that police and prosecutors may look into, thus perhaps reducing the risk of their abusing their discretionary powers.¹⁶

This argument also has drawbacks. First, while it may give reason for the state not to punish harmless risky conduct, it is little help with attempts. It does not explain why the state might have reason to punish attempts less harshly than would be permitted on culpability grounds (i.e., as harshly as completions). Since attempt is already a crime, and merely tends to be punished less harshly, police and prosecutors would still be authorized to pursue attempts. Thus, it is not clear how the present idea explains why Bert is punished less than Ernie.

A second problem is that withholding some punishment when harm is lacking is not the only way the legislature can combat abuses by law enforcement. Instead, the legislature might fully criminalize all the conduct it in principle may on culpability grounds, but then impose heightened accountability to discourage abuses by law enforcement. Perhaps police would face summary dismissal for using overly aggressive tactics. Perhaps prosecutors could be ordered not to always charge the maximum they can plausibly justify (as former attorney general Jeff Sessions ordered federal prosecutors to do), but rather to charge less serious crimes at times (as was the policy under another former attorney general).¹⁷ Why would this alternative way to prevent abuses be worse than punishing Bert and Jack less harshly than Ernie and Jill, respectively? It is an empirical question whether this would be a more effective solution, but it is conceivable it could work better, as it directly targets the abuses to be prevented.

Thus, while punishing harmless actors less than their harmful counterparts is one possible way to combat abuses, this punishment differential is not a *narrowly tailored* way to combat law enforcement abuses. Thus, I doubt it is on its own adequate to justify the many ways the law actually takes harm as the basis

16 Admittedly, it is an empirical question whether restricting the crimes on the books lowers the rate of law enforcement abuses. It is conceivable that reducing the opportunity to abuse one's discretionary powers with respect to some crimes would simply shift the abuses over to other crimes.

17 Compare Attorney General Jefferson Sessions III, "Memorandum for All Federal Prosecutors on Department Charging and Sentencing Policy" (May 10, 2017) with Attorney General Eric Holder Jr., "Memorandum for All Federal Prosecutors on Department Charging and Sentencing Policy" (May 19, 2010).

for differential punishment. It does not fully solve the legal luck problem.¹⁸ Instead, we need a reason for ratcheting down punishment for luckily less harmful actors that applies even with perfect investigation and enforcement mechanisms in place.

2. A NEW RATCHETING-DOWN JUSTIFICATION

I now want to develop a new ratcheting-down argument that fares better than the previous versions. I do not contend it is the only argument that can succeed. Core features of criminal law doctrine—like the tendency to punish harmless wrongdoers less harshly—need not rest on one single justificatory argument but could have multiple normative foundations. Still, the argument I develop has distinct advantages. It is the right kind of argument (i.e., a normative, fairness-oriented reason, not a self-interested consideration). It covers the full range of cases to be explained and does not rely on highly contingent facts (just more stable background conditions). And unlike the argument from preventing law enforcement abuses, it is *narrowly tailored* in the sense of showing why withholding some punishment from harmless actors satisfies the relevant evaluative commitments *directly*. The argument would apply even if our general law enforcement mechanisms were perfected.

As we will see, one might think the normative commitments behind the argument apply more forcefully to endangerment than attempts. Thus, while the argument provides a general recipe for how to defend luck in criminal law doctrine, one need not use it to justify all forms of luck in the criminal law. I show how the argument can be extended to cover attempts as well as endangerments, but I am also happy if it proves best to use the argument only piecemeal for some crimes rather than as a general defense of legal luck.

The argument proceeds in two stages. First, I argue that the legislature has a *pro tanto* duty not to be cruel in passing criminal laws. Second, I argue that withholding some punishment from harmless actors is a particularly appropriate strategy for avoiding an important sort of breach of this duty. The upshot is a justification for the legislature to ratchet down punishments imposed on harmless wrongdoers. I do not claim that this policy is the *only* defensible way a legislature can satisfy the normative principles behind the argument, but it

¹⁸ There may be other institutional ratcheting-down arguments. Perhaps withholding some punishment from nonharmful actors would help reduce the costs of incarceration. However, this idea has similar drawbacks. It is an empirical question whether this would be an effective way to achieve the desired cost reductions. Again, I am open to cost reduction as an additional benefit of recognizing luck in the law, but I doubt it is sufficient by itself to provide the sought-after justification.

provides one sufficient normative ground for the legislature to withhold some punishment that harmless wrongdoers otherwise deserve.¹⁹

2.1. *The Legislature's Duty Not to Be Cruel*

My argument starts from the idea that the legislature has a *pro tanto* duty not to be cruel. By this, I do not just mean cruelty in the narrow sense of taking pleasure in another's suffering, which is an extreme instance of cruelty. The concept of cruelty also covers less extreme cases.²⁰ One can be cruel through callousness and insensitivity to harm.²¹ Conduct can be cruel if it displays a notable lack of sympathy or concern—as when one is cold, indifferent, or overly rigid in one's treatment of others. Beyond this, I remain neutral on precisely what cruelty consists in. For present purposes, I do not need a full account of the notion, as I rely only on minimal claims that can be widely accepted. In general, though, I contend that this duty applies also to the legislature. When passing criminal laws, the legislature has a *pro tanto* duty to eschew cruelty in its treatment of those subject to its laws.²²

The “*pro tanto*” qualification indicates that other considerations can sometimes outweigh this duty. Thus, one might be tempted to say instead that the legislature merely has weighty moral *reasons* not to be cruel or callous when legislating within the criminal law. This is true but too weak. Although it is overridable, we are still dealing with a *duty* (albeit a *pro tanto* one) because,

- 19 I do not need to show that recognizing moral luck is the *uniquely best* way to satisfy the normative principles underlying this argument. A legislature is generally permitted to make incremental progress on a given challenge and need not always adopt the best possible solution to that problem (especially where the ideal solution carries other costs or is difficult to implement).
- 20 See Barrozo, who distinguishes four conceptions of cruelty, including agent-focused vs. victim-focused notions (“Punishing Cruelly,” 69–70).
- 21 For example, even if those responsible for factory farming do not take pleasure in the animal suffering caused, some of these practices can still be cruel if they manifest callousness by failing to minimize animal suffering as much as possible (even at considerable cost).
- 22 I assume that this notion of legislative duties is sensible and am happy to employ whatever proves to be the best account of it. One might be a reductionist who takes our talk of legislative duties to be shorthand for the duties of individual legislators. Or one might adopt a nonreductionist view. Nonreductionism seems apt for the legislature given that its sophisticated decision-making capacities may be enough to hold it responsible and regard it as bearing duties in its own right. See, e.g., List and Pettit, *Group Agency*, 153–63. Nonreductionism also makes sense for legislatures as they can adopt positions—through negotiations or premise-by-premise decision-making—that are independent of the personal views of individual legislators and which no individual legislator fully endorses. See Lackey, “Group Knowledge Attributions.” The present argument does not require nonreductionism, however.

all else equal, a complaint—even blame—would be legitimately leveled by citizens at the legislature if it unjustifiably breaches its duty to eschew cruelty.

Thus, the *pro tanto* duty to not be cruel is stronger than the reasons to display a virtue like mercy. Mercy, we saw above, is generally thought to be optional—something it might be good for the legislature to display in action, but not something it *must* display. By contrast, the *pro tanto* duty to not be cruel is not similarly optional. Rather, it is something the legislature must do on pain of being an appropriate target of complaints and blame all else equal. This is why the present argument is supposed to be stronger than the argument from mercy.²³ It is not plausible to say that there is a duty to show mercy (or that people have a right thereto), but we do plausibly have a right for our institutions not to be cruel to us.²⁴ Behaving cruelly would open up these institutions to legitimate complaints—even blame. This duty can be summarized thus:

Legislative Duty to Not Be Cruel: When legislating in the domain of criminal law, the legislature has a *pro tanto* duty not to be cruel—that is, weighty moral reasons not to be callous, mean, vicious, or vindictive toward the affected citizens—with the result that if these reasons are not outweighed or defeated by other reasons, a legitimate complaint, and plausibly some blame, would be fittingly directed at the legislature in virtue of this failure.

I am not wedded to the details of this statement of the duty. I am happy to use other concepts or formulations if preferable. Nor do I claim that the legislature is the only government actor to have such a duty.²⁵

23 Note also that the duty to not be cruel is a stronger imperative than one's general reasons not to display vices. For at least some vices—especially cowardice or stupidity—there are weighty reasons not to display them in conduct, but doing so nonetheless is not a failing for which blame or complaint is appropriate. There are weighty reasons against being cowardly and stupid, but since such behavior need not manifest insufficient regard for others, it does not automatically call for a blaming response. By contrast, the duty to not be cruel is stronger in that being cruel in one's actions *does* manifest insufficient regard, and thus merits a blaming or complaining response (at least assuming there are no countervailing considerations that justify the failing). Thus, being cruel is more like the vice of being unjust, which also would call for a blame response or generate a complaint. Hence, my claim is not simply reducible to the truism that we generally have reasons not to display vices in our conduct. See also section 3.2.

24 For a discussion suggesting that there is no duty to show mercy, see Tasioulas, "Mercy," sec. 4. Similarly, no complaint or blame becomes warranted if one consistently declines to show mercy—if one in this sense is "merciless."

25 I am open to thinking that other official actors—including courts, prosecutors, and even the executive—may also have an analogous duty to not be cruel within their areas of responsibility.

I cannot give a full account of everything needed to avoid breaching this duty. Articulating *ex ante* all the ways one might be cruel, callous, vicious, or vindictive will be extremely difficult, given how nuanced and context sensitive these notions are.²⁶ Instead, all I need is something more modest: a sufficient condition for being cruel.

More specifically, all I need for the present argument is the claim that *one* way for the legislature to breach its duty to not be cruel would be for it to always impose the full amount of criminal liability on citizens that would be permissible on culpability grounds. It would show the legislature to be cruel if the legislature were to always impose the maximum penalties it could get away with given the culpability of the actor—that is, the maximum penalties that would not run afoul of the desert constraint, which prohibits disproportionately harsh punishments.²⁷ Even if always imposing the maximum criminal liability permitted would not violate any deontological side constraints that restrict the criminal law, systematically imposing the harshest permitted punishments would reflect badly on the legislature. It would suggest the legislature is out to always extract its pound of flesh, no matter the cost and no matter how much insensitivity and callousness it manifests. We can succinctly put the point as follows:

Sufficient Condition: If the legislature, when passing criminal laws, always imposes the maximum amount of criminal liability on citizens that would be permitted on culpability grounds (i.e., as much liability as it can without offending the proportionality constraint built into negative retributivism), then this shows the legislature to be callous, mean, and vindictive—that is, cruel—at least provided there are no sufficiently weighty countervailing reasons or defeaters that would justify or excuse this conduct by the legislature (such as an abnormally acute need for heightened deterrence).

Thus, the legislature would open itself up to complaints—even blame—if it always imposed the maximum punishment permitted on culpability grounds.²⁸

26 To be clear, by “vicious,” I do not mean the broad idea of displaying vices in one’s conduct—as one might think based on note 23. Instead, I intend a more colloquial meaning of “vicious” (i.e., being nasty and hostile).

27 See, e.g., Berman, “The Justification of Punishment,” 151.

28 For example, former US attorney general Jeff Sessions’s prosecutorial policy, which ordered federal prosecutors to always charge the maximum crimes that could be maintained, was decried as cruel. The director of Human Rights Watch argued that this policy “ignor[ed] the facts about the cruelty, waste, and ineffectiveness of the ‘tough on crime’ policies of the 1980s and 90s” (Jackman, “Sessions Takes Federal Crime Policy Back to the ‘80s”).

What grounds this claim? Why would it be cruel to always impose the maximum punishment allowed on culpability grounds? Let me offer some substantive considerations in favor of this claim. As we will see, these considerations may carry more weight as applied to risk-taking than to acts done with intent to cause harm.

2.1.1. *Why Ratchet Down for Risky Conduct*

Most importantly, the above view derives support from the simple realization that people are imperfect, and the legislature may be criticized for not attaching due weight to this fact. Granted, at each moment, it is plausible that each of us ought not behave in ways that are unduly dangerous, risky, or otherwise wrongful. This is a *synchronic duty*. Nonetheless, it would be unreasonably demanding to expect everyone to act perfectly safely and reasonably at every moment across a long period of time—to attain *diachronic perfection* across a period of years. For practically everyone, some moral failures—at least minor ones—are eventually unavoidable. There will be times when all of us behave in ways that are dangerous enough to make us fitting targets of criticism and blame. Perhaps this will be due to tiredness, distraction, cognitive failings, stress, or other excusable burdens. We are likely to face provocations, frustrations, and stresses that over time accumulate in ways that naturally lessen any normal person's capacity for self-restraint (particularly if compounded by nonideal cognitive or emotional conditions). Without pretending to a level of precision I cannot obtain, I submit that at a minimum, a few culpable screwups every few years are inevitable. Thus, even if there is a synchronic duty, applicable at each point in time, not to act in ways that are unduly risky, and which we can be fittingly blamed for breaching, it would be overly harsh for the criminal law to demand unassailable behavior at every moment across long periods (like a term of years). That is, the criminal law should not insist on diachronic perfection across long stretches of time. This is particularly true where the state itself bears significant responsibility for creating trying and burdensome conditions that make it more difficult to exercise the restraint and care necessary to attain diachronic perfection across long periods. To insist on diachronic perfection across long stretches of time would be cruel given how far beyond the actual capabilities of most normal people it is to attain such levels of perfection (at least without entirely sacrificing many valuable activities we should be free to pursue). Thus, it would be harsh and unreasonable for the criminal law not to make accommodations for this fact.

Therefore, even if it would be permitted on culpability grounds to impose punishment whenever we do something unduly risky—which amounts to insisting on diachronic perfection—doing so would make the criminal law

more vindictive and cruel. Since very few would realistically be able to attain diachronic perfection, especially given a sufficiently long time frame, a legislature that always demanded it in its criminal laws would be cruel—even if this would also be permitted on culpability grounds.

The upshot is that in light of normal human imperfection, the legislature has weighty reasons to find ways not to always punish to the maximum extent permitted in response to what we are fittingly blamed for synchronically. To avoid being cruel, the legislature is morally required to pass criminal laws that make some meaningful concessions to unavoidable human imperfection. I contend that all else equal, we as citizens have a *complaint* against the legislature if it fails to make such concessions and instead insists on diachronic perfection by always imposing the maximum punishment permitted on culpability grounds. This supports the above view, as stated in Sufficient Condition.

2.1.2. *Why Ratchet Down for Intentional Harm*

Perhaps this thinking has limits. While some relaxations of criminal liability may be warranted for risky conduct because people are neither perfect nor perfectly good, this reasoning may seem less plausible for graver intentional wrongs like murder or theft. It is far from inevitable that most of us will commit such serious wrongs in our lifetimes. Thus, might the argument only support withholding some punishment for lesser categories of misconduct?²⁹

Without conclusively settling the question, note that the present reasoning still has some merit for actions (like attempts) done with intent to harm—for two reasons. First, even the best of us can be unfortunate in the circumstances we face. Even good people can encounter serious provocations and trying circumstances, which require great effort and restraint not to succumb to. Perhaps in ideal conditions—when well rested, well fed, well paid, and well supported emotionally—good people will always manage to resist provocations or temptations and stay on the right side of the law. But as trying circumstances add up over time, as conditions become less ideal, and as we extend the time frame, maintaining diachronic perfection becomes less likely—even for the otherwise virtuous. Thus, the argument plausibly does support withholding some punishment even for *some* serious wrongs as a concession to human imperfection.

Second, there can be extenuating circumstances for some kinds of property crimes and perhaps even some acts of violence, which should not be recognized as formal defenses—as Judge Bazelon’s proposed “rotten social background” defense would have been—but which nonetheless put normative pressure

29 Thanks to Erik Encarnacion and Liat Levanon for helpful discussions on this point.

on the legislature to make concessions when passing criminal laws.³⁰ Especially when the legislature itself bears some responsibility for allowing severe inequality and other criminogenic conditions to persist, it may be cruel for the legislature to always insist on the maximum punishment permissibly imposed on culpability grounds. The suggestion is not that this means a new type of defense must be recognized. Rather, it is another reason for the legislature to seek ways to relax punishments in order to avoid being cruel—and this reason would also cover intentional crimes.

If one does not find these considerations compelling, then this would mean one takes the legislature not to have as weighty a duty to find ways to ratchet down punishments for intentional crimes. This, as seen below in section 2.3, may be what explains why punishment is not *fully* withheld from attempts but is only imposed at a reduced rate. Regardless, we have seen some considerations that may help underwrite the legislature's duty to not be cruel also where intentional crimes are concerned.

2.2. *Withholding Some Punishment from Harmless Wrongdoers Is an Apt Way to Avoid Being Cruel*

I have argued that the legislature, when passing criminal laws, has a *pro tanto* duty to avoid cruelty—including cruelty of the sort Sufficient Condition specifies. The criminal law must make *some* concessions to normal human imperfection (or other considerations supporting relaxation of punishments) by not insisting on diachronic perfection across long stretches of time. Instead, it must somehow ratchet down punishments below the maximum permitted on culpability grounds.

However, the legislature has great flexibility in deciding how to discharge this duty. There is no limit to the ways it could avoid being cruel in the way Sufficient Condition specifies. All it would have to do is find *some* meaningful way not to impose the maximum punishment permitted on culpability grounds. In this respect, the requirement to avoid cruelty of the sort Sufficient Condition specifies is like the requirement to give to charity. We have no duty to give to any particular charity, but we do typically have a duty to give to some charities sometimes. Otherwise, we would show ourselves to be callous and unkind (all else equal). But we have wide discretion in how to avoid being-callous-by-giving-to-no-charities.

The question thus is *how*, exactly, the legislature should discharge its duty to not be cruel in the way Sufficient Condition specifies. The legislature needs a

30 For discussion of issues relating to the “rotten social background” defense, see, e.g., Morse, “The Twilight of Welfare Criminology,” 1252.

way to break the impasse and decide how best to refrain from always punishing as harshly as it in principle could. I argue that withholding some punishment from harmless wrongdoers is an especially appropriate way—better than the natural alternatives—to avoid being cruel in the manner Sufficient Condition specifies. Defending this claim will complete the argument that the criminal law is justified in recognizing some luck. (To see the structure of the argument, focus for now on luckily harmless misconduct *in general*. I return to the differences between endangerment and attempts in section 2.3.)

Why, then, is the absence of harm a basis for withholding some punishment that is otherwise due harmless wrongdoers? Why is this a good way for the legislature to comply with its duty to not be cruel? Why not just flip a coin for each culpable offense to decide if some punishment should be withheld?

The answer is that the presence of clearly identifiable victims makes a normative difference—at least enough of a difference to break the impasse the legislature faces in deciding how to comply with its duty to avoid cruelty. Unlike other theorists, I do not go so far as to claim that the presence of victims who are harmed is itself a reason to impose criminal liability, or would otherwise (by itself) support enhancing punishments.³¹ Rather, I rely only on the more modest, and hopefully less controversial, claim that the presence of victims is a consideration that can tip the scales in favor of one way of complying with the legislature's independent duties rather than other ways.

To see why victims can make a difference to how the legislature should avoid being cruel, distinguish cases where culpable conduct causes harm from those where it does not. For the former cases, most will include distinct and identifiable victims—whether someone who was directly harmed or their nearest and dearest. Where distinct and identifiable victims exist, they will have a claim on the state to acknowledge their rights that were violated—more precisely, to reaffirm these rights by seeing to it that justice is done to the relevant wrongdoers. In such cases, the victims are in a position to loudly and intensely press these claims against the state, and they can be expected to do so if practically able. This is because when harm occurs, given human psychology, the danger

31 Cf. Binder, “Victims and the Significance of Causing Harm.” As Binder puts it:

What I am suggesting is that we punish harm not only in order to express something to the offender and about the offender, but also to express something to the victim and about the victim to others. We punish not only in order to admonish the offender that he or she should respect the victim, but also in order to show the victim our own respect. If so, we are punishing harm for a purpose that transcends doing justice to the offender. (“Victims and the Significance of Causing Harm,” 733)

See also *infra* note 35 and accompanying text.

will tend to seem more salient, and so victims are more likely to experience resulting feelings of anxiety and insecurity, all else equal.³² This fact plausibly grounds a legitimate claim against the state for it to relieve these painful feelings by providing reassurance to victims through reaffirming their rights.³³

- 32 This is an empirical claim, though I think it is *prima facie* plausible. There is research supporting it. One large study shows higher rates of serious psychological effects (suicide attempts, suicidal ideation, “nervous breakdowns”) in victims of some serious crimes—such as rape—compared to victims of analogous attempts. See Kilpatrick et al., “Mental Health Correlates of Criminal Victimization,” 869–70. Nonetheless, the evidence remains mixed regarding attempts (while not speaking to endangerments). Specifically, the higher rates of serious psychological effects were *not* seen in connection with other types of crime, such as completed vs. attempted robbery or completed vs. attempted molestation (869–70). Here, the attempts actually carried higher rates of serious psychological effects. The authors explain as follows:

Whereas completed rape was much worse [psychologically] than attempted rape, attempted molestation and attempted robbery had more negative mental health consequences than did their completed counterparts. This finding is counterintuitive but may be partially explained by the observation that attempted attacks leave much room for ambiguity in the victim’s mind as to what the assailant intended and as to the actual danger the victim was in. The extent to which victims in ambiguous situations attribute very dangerous intentions to their assailants is apparent in the finding that 35% of these victims of attempted molestation, compared with 18% of victims of completed molestation, thought they were likely to be killed or seriously injured during their assault. Victims of attacks that were not completed do not know what they escaped. (872)

Note that these findings apply to attempts only and do not undermine the plausibility of the analogous claim about endangerments (namely, that suffering a given personal harm tends to be psychologically worse than merely being subject to a risk thereof). Moreover, regarding attempts, even the mixed results above are still compatible with the claim that being the victim of a completed crime on the whole tends to be psychologically worse than being the victim of the analogous attempt at least *all else equal*—including knowledge of the perpetrator’s intentions and the danger the victim was in. This would still provide some support for my argument, which is concerned with assessing the normative strength of competing claims by personally harmed victims vs. those who were in danger. Nonetheless, it remains true that if no version of the empirical claims my argument requires are supported by the evidence, the argument would fail. With this important caveat, I proceed under the assumption that at least some qualified versions of the empirical claims I need are plausible enough to warrant considering normative arguments based on them.

- 33 Of course, it will not always be the case that there will be anyone practically able to press such a claim. For example, the victim might be unable to press the claim because she was killed and no one can do it for her. (Perhaps she was a hermit with no friends or family.) Still, all I claim is that when criminal wrongdoing causes personal harm to a distinct victim, then a claim arises that this person, or someone acting on her behalf, is *entitled* to press. The victim, or someone close to her, will be likely to do so if practically able. That is enough to get my argument off the ground.

The other kind of case is where the wrongdoer's conduct does not cause direct harm—perhaps merely due to luck, as with Bert (the lucky attempter) and Jack (the lucky endangerer). In these cases, there typically will not be victims who are as distinct and identifiable as in the former kind of case, where harm ensues. When harm does not occur in ways that produce distinct and identifiable victims (whose rights were violated in an especially salient way), there will not be anyone with an equally strong claim against the state to have their rights reaffirmed through the imposition of criminal liability. Less anxiety and distress are likely to be felt, all else equal; and with less insecurity to be alleviated, there is a less weighty claim against the state to reaffirm the rights violated when no harm occurs.³⁴

To this, one might object that being exposed to undue risk or targeted in an attempt, even if not directly harmful to body or property, could still violate a right, generate distress and anxiety, and thus generate a legitimate claim for the reaffirmation of rights. Would not people whose rights are violated non-harmfully also want the state to acknowledge these rights through a criminal law response?

Yes, but in the main, these claims for the reaffirmation of rights will not be as intensely felt, stem from as much anxiety or distress, and thus be likely to be pressed as loudly as the analogous claims by victims whose rights were harmfully violated—such as by a punch, physical wound, or psychological trauma. Given our psychology, harmful rights violations are likely to cause more anxiety and distress, and thus generate more pressing and loud claims for the state to reaffirm the violated rights than would be expected for analogous nonharmful rights violations. Indeed, when rights violated harmfully are not reaffirmed by the state, this is likely to leave in place more fear, insecurity, and distress for victims than when rights violated *non-harmfully* are not reaffirmed. Thus, the legislature would be justified in ascribing more weight to claims for the reaffirmation of harmfully violated rights than claims stemming from harmless rights violations—or at least this is so when the legislature cannot respect *all* such claims but must choose between them. And in this context, the legislature *must* choose because of its duty not to be cruel.

Granted, if the state did not have to choose which claims for reaffirmation of rights to respect, then it could just respect them all by imposing criminal liability in response to any serious rights violation—whether harmful or not. But in this context, the state *must* choose which of these claims to respect because of the legislature's duty not to be cruel in the way Sufficient Condition specifies—that is, not to display the cruelty of always punishing the maximum permitted on

34 See *supra* note 32.

culpability grounds. Since the state must withhold full punishment sometimes, it cannot respect all claims to reaffirmation of rights that could be pressed in any case of culpable wrongdoing. Thus, the state must choose. And in so doing, it would be reasonable to ascribe less weight to claims for reaffirmation of rights stemming from nonharmful rights violations than to similar claims stemming from harmful rights violations.

This shows how the presence of harm can make a difference. Suppose I am right that harmful wrongdoing (like Ernie's and Jill's) typically generates claims to reaffirmation of rights that are louder, more pressing, and based on greater anxiety—claims to which the legislature can legitimately ascribe more weight—than the comparable claims arising from harmless wrongdoing (like Bert's and Jack's). If so, harm can make enough of a normative difference to help the legislature decide when it should withhold full punishment—as it must do in some way or else be criticized as cruel and callous. If the legislature withholds full punishment from some cases where harm ensues, it will be disregarding the weightier claims by distinct and identifiable victims to have their violated rights reaffirmed. Call these the *loud claims for reaffirmation of rights*. But if the legislature withholds full punishment only from cases where harm does *not* ensue, it will not be disregarding any of these loud claims for reaffirmation that arise when harm ensues. Assuming there *are* claims for reaffirmation of rights when the wrongdoing is harmless, these claims would typically be ones with less weight—ones that would not be pressed as loudly and intensely, as they would tend not to stem from as intense anxiety. Call these the *quiet claims for reaffirmation of rights*.

Since the legislature must sometimes withhold punishment to avoid being cruel, it would be justified in deciding to do this in a way that at most disregards only the *quiet* claims for reaffirmation of rights (arising when there is no harm), but that always fully satisfies the *loud* claims for reaffirmation of rights (arising when harm ensues). This approach would be a less worrisome failure to reaffirm victims' rights in general. The legislature thus would be justified in seeking to avoid being cruel in a way that does not disregard the loud claims. If some claims to reaffirmation of rights must be disregarded, better to disregard the quiet kind that may arise from harmless wrongdoing than the loud kind arising from harmful wrongdoing.

Before considering differences between reckless endangerment and attempts, let me clarify a central point. What is the role of the duty to not be cruel in the argument? Why not just appeal to victims' rights directly? The reason is that I want my argument to be as ecumenical as possible. I do not want to rely on the contestable view, which others have recently used to argue for luck in the criminal law, that the presence of harmed victims—that is, the need

to express respect for them and reaffirm their rights—can itself be a sufficient basis for imposing or enhancing punishment.³⁵ Whether this view is correct is a fraught question, one I do not want my argument to hinge on. Accordingly, I do not maintain that victims' claims for the reaffirmation of rights have any independent (nonderivative) justificatory force that militates *in favor* of punishment. Instead, I maintain only that such claims of harmed victims bear on when to *ratchet down* punishment, and I only take them to have this normative significance *in virtue of* the more fundamental legislative duty to not be cruel in its criminal laws. This duty requires the legislature to choose some violated rights that will remain unreaffirmed. On my view, the claims of harmed victims for reaffirmation of rights only get to serve as a tiebreaker on this issue because the legislature must make this choice. Hence, my rationale for withholding some punishment from harmless wrongdoers is meant to be more modest and more widely acceptable than more ambitious victim-focused arguments recently offered in favor of luck in the law.

2.3. *Does the Argument Apply across the Board?*

My argument provides a recipe for how to justify withholding full criminal liability from some categories of culpable but harmless wrongs. As indicated in section 2.1, some might think the argument is less plausible for attempts, as

35 See Boeglin and Shapiro, "A Theory of Differential Punishment." They defend a victim-facing justification "premised on the notion that the state should take the interests of victims into account when determining how severely" to punish, and they contend that in at least some instances, greater punishment can justifiably be imposed on harmful actors "out of respect" for victims of the harm caused (1503). See also Binder, "Victims and the Significance of Causing Harm."

My argument here is different from Boeglin and Shapiro's not only because I offer a "ratcheting-down" argument, while they argue for "ratcheting up" punishments for harmful offenders. They, for example, express sympathy for the "judgment that, at times, the degree of punishment warranted by offender-facing justifications might seem 'insufficient' in light of the harm that a victim has suffered" (1523). More importantly, as noted, my argument strives to be more ecumenical. The victim-facing considerations I adduce merely serve as a reasonable way for the legislature to decide how to satisfy its duty to not be cruel. Unlike Boeglin and Shapiro, I do not contend that victim-facing considerations by themselves can justify a policy of differential punishment as between harmful and harmless wrongdoers. Instead, my argument gives victim-facing considerations a role to play in justifying moral luck only *in virtue of* the duty to avoid cruelty and viciousness in the criminal law. Without the normative force of this duty to avoid cruelty, I doubt that victim-facing considerations alone suffice to justify imposing more punishment on harmful wrongdoers than on their harmless counterparts.

intentional wrongs, than for endangerment.³⁶ Nonetheless, the argument need not be rejected outright for attempts. It can still apply in modified form.

2.3.1. *Endangerment*

Begin with the argument in its pure form as applied to endangerment. The duty not to be cruel requires the legislature to find meaningful ways to impose less criminal liability than would be permitted on culpability grounds. In deciding how, it may look to the normative difference victims make. Consider three policies the legislature may adopt to avoid being cruel in the way Sufficient Condition specifies:

1. Unduly risky conduct is a crime only when harmful, not when harmless.
2. Unduly risky conduct is a crime only when harmless, not when harmful.
3. A coin is flipped in each case of unduly risky conduct, regardless of harm, to determine whether to impose less punishment.

Policy 2 is least justified. In virtually all cases of risky conduct that causes harm, there will be distinct and identifiable victims who would have loud and weighty claims for reaffirmation of their rights. But all these loud claims would be disregarded by 2. When wrongdoers cause no harm, claims for reaffirmation of rights are less likely to arise—and if they do, it would only be the quiet kind to which the legislature would be justified in attaching less weight (since they are likely to stem from a less intense sense of insecurity). Policy 2 disregards all the loud, weighty claims to reaffirmation of rights, while only respecting the quiet ones. Policy 3 fares better but remains suboptimal. Under this policy, the state would end up disregarding loud claims for reaffirmation of rights in half the cases where such claims arise.

Policy 1 does the best job of respecting claims for reaffirmation of rights, while also complying with the legislature's duty to not be cruel. When risky conduct causes no harm, there are no distinct and identifiable victims who have loud, weighty claims for reaffirmation of rights. Admittedly, by not punishing harmless risky conduct, the legislature may fail to respect some claims of the quiet kind. Even if quiet claims do arise when no harm results, they are unlikely to be as intense and pressing as loud claims, all else equal, and so the legislature would reasonably attach less weight to them than the comparable loud claims. Thus, a legislature could reasonably conclude a good way to comply with its duty to not be cruel is to refrain from criminalizing risky conduct except when it causes harm. This would ratchet down some punishments while fulfilling the

36 See, e.g., American Law Institute, *Model Penal Code*, sec. 5.01, which holds that attempts require intent, not mere recklessness.

weightiest claims to reaffirmation of rights and only disregarding such claims when less weighty.³⁷

2.3.2. *Attempts*

Some might think this argument is less compelling for attempts, as intentional wrongs. After all, failing to withhold some criminal liability from luckily harmless attempters might seem not to make the legislature *as cruel* as failing to do so for harmless endangerers. As noted in section 2.1, it seems harsher to insist on long periods of diachronic perfection in avoiding risky conduct than to insist on such diachronic perfection in avoiding intentional wrongs. If avoiding cruelty in general requires making some concessions to natural human imperfection, the legislature might seem *more cruel* in failing to allow for natural imperfection with respect to risk-taking than where intentional wrongs are concerned. Thus, the pure form of the argument might not be as compelling for attempts, as intentional wrongs. Of course, some may think concessions to human imperfection are due even concerning intentional wrongs (perhaps particularly if the legislature bears responsibility for creating criminogenic conditions). But even for those who find the argument less forceful for intentional wrongs, it need not be rejected outright. It might still apply to attempts in modified form.

To see this, note that the law does not *completely* withhold punishment for attempts. It merely imposes *less* criminal liability on Bert, for example, while imposing full liability for Ernie's equally culpable harmful conduct. The present rationale can explain why.

It still seems overly harsh of the legislature not to ratchet down *at all* for intentional wrongs as a concession to human imperfection, but the need to do so is less pressing. Instead, the legislature might calculate that fully withholding punishment from harmless attempters would be too unfair to victims. It might reasonably determine it had better *at least somewhat* reaffirm the violated rights of those who were the target of an intentional wrong—even when no harm results. After all, attempting a crime requires intending it. Thus, when one is the target of an attempt, one's rights have been violated in a more salient way—one that is more serious, all else equal, than when one was merely subject to undue risk of the analogous harm but was not targeted in an attempt. Thus, in attempt cases, there are more likely to be distinct and identifiable victims with weightier claims for the reaffirmation of rights—though not as weighty as these claims would be if harm ensued. Accordingly, the legislature would legitimately feel pressure to provide *some* criminal law response to the claims for reaffirmation

37 The same story could be used to justify punishing criminal negligence only when it causes harm as well.

of rights by victims of attempts. However, since the legislature *also* must find meaningful ways to withhold full punishment in order to eschew cruelty, it would have reason to seek a balance between the competing normative pressures it faces. It must not only (a) respect the still somewhat weighty claims of victims of attempts for the reaffirmation of rights but also (b) balance this against the need to sometimes withhold full punishment to avoid being cruel. How to strike this balance? A plausible answer is to punish attempted crimes less harshly than analogous completions.

Thus, a modified version of the argument applies to attempts. Consider three policies:

4. Punish mere attempts less harshly than the analogous completed crimes.
5. Punish completed crimes less harshly than mere attempts.
6. Flip a coin in each case of a completion or attempt to decide whether to punish the conduct at a reduced rate.

Policies 5 and 6 disregard victims' claims to the reaffirmation of rights to a greater degree than policy 4. Policy 5 gets it exactly backward in partially frustrating the loud claims of harmed victims while fully satisfying only the less weighty claims of victims of mere attempts. Policy 4 gets it the right way around. It fully satisfies the loud claims of harmed victims and only partially frustrates the less weighty claims of attempt victims. While 6 fares better than 5, it still does not do as good a job as 4, which is the most justified of the trio.³⁸

Thus, even if one finds it less imperative for the legislature to withhold some criminal liability for intentional wrongs than for endangerment, this does not mean one must entirely reject the ratcheting-down argument for attempts. One could simply adopt the modified version outlined. This would explain the prevailing legal practice of punishing attempts, although less harshly than analogous completions. I remain neutral on whether ratcheting down for attempts is truly needed for the legislature to satisfy its duty to not be cruel. But the reasoning has plausibility even here.

38 A legislature might have adopted the attempts solution for reckless endangerment. Rather than withholding punishment from harmless reckless conduct altogether, the legislature might have decided to withhold only *some* of the punishment that is due. The legislature might think a policy like 4 is normatively better for reckless endangerment than policy 1. I will not try to resolve whether 1 or 4 for reckless endangerment better accommodates both the need to respect victims' claims for reaffirmation of rights and the legislature's duty to not be cruel. Which policy is better may depend on contingent facts about attitudes in the jurisdiction. Even if reasonable legislatures differ on this point, my primary aim here is just to illustrate the *kind* of reasoning that would justify some policy like 1 or 4 that recognizes luck.

3. ALTERNATIVES AND OBJECTIONS

3.1. *Alternative Solutions*

To show that recognizing luck is an appropriate way for the legislature to avoid being cruel, I need to explain why this legislative approach is no worse—and preferably better—than the natural alternatives.³⁹ I cannot canvass all alternatives. But by showing why recognizing luck is not clearly worse than the obvious alternatives, I aim to demonstrate that there is at least an *adequate* normative justification for recognizing luck in the criminal law—that it stands as an available option for the legislature.⁴⁰ I consider four other ways to ratchet down punishments below what is warranted on culpability grounds. I am in no way opposed to these measures, but I claim that they do not obviously do a better job than recognizing luck does in satisfying the considerations behind the duty to not be cruel.

Consider the first two alternatives together. They involve relying on the discretion of other state actors—either prosecutors (or other law enforcement officials) or sentencing judges—to ratchet down the punishments otherwise due on culpability grounds. This move is not satisfying for several reasons. First, prosecutorial and sentencing discretion are worrisome insofar as they are *ex post solutions* to the problem of a legislature passing laws that count as cruel. Better for the legislature to prevent this problem from arising *ex ante* by passing laws that avoid the issue. Second, there are concerns about the legislature delegating its responsibility to avoid cruelty. If the legislature leaves ratcheting down to other actors' discretion, it cannot be as confident that the required ratcheting down will actually happen. There is no guarantee that prosecutors or judges will ratchet down as needed to satisfy the legislature's duty to not be cruel. Safer for the legislature itself to see to the satisfaction of this duty.

Most importantly, the legislature's duty to not be cruel places constraints on the *content* of the criminal law, not merely its enforcement. Even if prosecutorial and sentencing discretion were used in a generous ratcheting-down manner, we would still have a complaint against the legislature for having put laws on the books that are cruel, callous, and vindictive in failing to make adequate accommodation for human imperfection and the legislature's perpetuation of criminogenic conditions (if applicable). To satisfy the legislature's duty to not

39 Many thanks to James Manwaring for pressing me on this point.

40 It is a familiar point that the legislature need not always adopt the best conceivable solution to a given problem but can be justified in taking steps that take us closer to the ideal solution than would otherwise be the case. It is in this sense that I aim to provide an *adequate* justification for recognizing moral luck, though not an argument that *mandates* it.

be cruel, such accommodation is needed in the *content* of the law—the substantive rules delineating criminal offenses.

A final worry about sentencing discretion in particular is that sentencing is individual focused, while the considerations behind the legislative duty to not be cruel are broadly applicable considerations involving pervasive human imperfection and (perhaps) criminogenic conditions sustained by the legislature. Sentencing judges typically respond to individual-specific factors affecting the defendant's conduct and circumstances. It would be unprincipled to announce that the defendant deserves a given sentence but then impose a lower sentence because of the general (non-individual-specific) difficulties in achieving diachronic perfection. Given the individual-focused nature of sentencing, judges are not in a good position to take account of the *generally applicable* reasons to make concessions to human imperfection. These are more properly the purview of legislatures.⁴¹

Similar considerations undermine a third natural alternative. Perhaps the legislature should accommodate the considerations behind the duty to not be cruel by expanding the affirmative defenses—especially excuses. This alternative may seem more promising because it involves the legislature itself making changes to the *content* of criminal law in order to ratchet down. Perhaps this could lead to new sympathy-oriented defenses like Judge Bazelon's "rotten social background defense," or a greater number of partial defenses that lessen the seriousness of one's offense.

Still, this alternative is not an optimal way to accomplish the required ratcheting down because excuses are individual specific. They are narrow individual-focused sets of conditions that call for a lower offense, or no conviction at all, for defendants who satisfy them. Determining whether the excuse is present requires looking at facts about the particular defendant—like whether he or she confronted especially challenging circumstances. The considerations behind the legislature's duty to not be cruel are broader, non-individual-specific facts about human imperfection, the inevitability of some culpable mistakes,

41 Here is a final reason sentencing judges sometimes cannot be relied on to fully satisfy the considerations behind the duty to not be cruel. Consider an offense with a mandatory minimum (and suppose that the mandatory minimum is not unjustly harsh). Now consider a defendant who is guilty of the crime but who deserves a punishment at the very bottom of the legally permitted sentencing range. The sentencing judge cannot lower this person's punishment any further without violating the mandatory minimum law. In such a case, the sentencing judge cannot satisfy the considerations that underlie the duty to not be cruel; only the legislature could do so by ratcheting down the whole permissible sentencing range, including the mandatory minimum provision setting the floor of the available punishments. Here, the legislature's duty to not be cruel could not even in principle be delegated to sentencing judges.

and (perhaps) the legislature's role in sustaining criminogenic conditions. Thus, individualized excuses are not an ideal vehicle for satisfying the more global considerations behind the duty to not be cruel.

Finally, a broader way for the legislature to eschew cruelty is to ratchet down punishments for all offenses.⁴² Jurisdictions like those in the United States arguably have weighty reasons to do this *anyway*.⁴³ But set that aside. Our question is this: Supposing that punishments are set at a nonexcessive level compared to culpability, what is a defensible way for the legislature to ratchet down punishments *even further* so as not to be cruel? Recognizing luck, I claim, is superior to lowering punishments across the board.

The reason is that luck does a better job on balance of satisfying victims' claims to reaffirmation of rights, thus alleviating the anxiety and insecurity grounding such claims. On the luck proposal, all *loud* claims possessed by harmed victims for the reaffirmation of rights will be fully satisfied, while only the quieter, less weighty such claims are not fully satisfied. By contrast, if we lower punishments for all offenses, *none* of these claims will be fully satisfied—neither the loud ones nor the quiet. Lowering the severity of punishments, say, 10–20 percent across the board would entail a corresponding degree of frustration of *all* claims to reaffirmation of rights. Given that *most* crimes involve harmed victims with loud claims to reaffirmation of rights, one can understand why a legislature might conclude that lowering all punishments does a worse job of *fully* satisfying people's claims to reaffirmation of rights, on the whole, than recognizing luck. The legal luck proposal, after all, always fully vindicates harmed victims' loud claims to reaffirmation of rights (which are due more weight). A legislature might reasonably conclude that, compared to the luck solution, ratcheting down across the board would leave in place more anxiety and insecurity on the part of victims.⁴⁴

I have not canvassed all alternatives to luck as a route to complying with the duty to not be cruel. However, I am not arguing that withholding some punishment from luckily harmless wrongdoers is the *uniquely best* way to avoid

42 While one might consider ratcheting down punishments only for *some* offenses, doing so requires a nonarbitrary way to decide *which* offenses this should be done for. That, however, is precisely the question that the occurrence of harm is supposed to answer. Ratcheting down the punishments of some offenses thus is not an alternative to the solution of recognizing luck—it is one attractive *instance* of this strategy.

43 Husak, *Overcriminalization*; Alexander, *The New Jim Crow*.

44 Furthermore, if the legislature ratcheted down across the board, it is likely that within a few years the population would become accustomed to the new range of punishments, so the lowered punishments would cease to be a salient way of not being cruel. By contrast, the luck proposal—given its differential treatment of equally culpable actors—is likely to remain a *visible* way of not being cruel even after a long time.

the form of legislative cruelty we are concerned with—only that this duty gives an adequate normative basis for recognizing some luck in criminal law.

3.2. *Objections*

My argument faces some objections that need a response.⁴⁵ First, is my reliance on the legislative duty to not be cruel just a disguised appeal to mercy? One might think there is a duty not to be merciless, which comes to much the same thing.

To start, I am not concerned with the terminology of the argument. In section 1.2, I rejected ratcheting-down arguments from mercy insofar as they conceive of mercy as optional. These arguments fail to the extent one thinks, as I do, that there is no duty to display the virtue of mercy. If one responds by moving to a stronger conception of mercy—one that sees mercilessness not merely as the absence of a desirable character trait but as the violation of a duty, which gives rise to complaints and blame—then the argument becomes quite similar to mine, albeit in different terminology. I think talking about a duty to not be cruel (or callous or mean, etc.) has greater force, and invites less confusion, since mercy may sound optional in a way that avoiding cruelty is not. But for those who prefer mercy talk, I say go for it. If what I have done is show how best to construct the argument from mercy, then that is progress too.

Still, substantive differences between my argument and the mercy argument remain. Mercy plausibly is individual specific. It is rendered sensible (nonarbitrary) in response to particular actions or feelings by the wrongdoer, such as apology, regret, repentance, or remorse. By contrast, as I have been at pains to argue (see section 2.1), the considerations underlying the legislative duty to not be cruel apply to persons *in general*—particularly the need to make accommodation for natural human imperfection, the practical unavailability of bad behavior especially over long periods of time, and the legislature's possible contribution to sustaining criminogenic social conditions. Thus, while mercy is based on specific features of the person or her behavior (things that make her merit mercy), the grounds of the duty to not be cruel do not turn on particular facts about people's lives or character. Hence, my argument remains different in substance from mercy-based arguments.

Here is a second worry. Suppose I am right that the duty to not be cruel requires the legislature to find ways to ratchet down some punishments. Where does it end? When have we done enough to satisfy this duty? Are endless relaxations of the criminal law required? No. As we ratchet down punishments more and more because of the duty to not be cruel, at some point the positive

45 Thanks to Erik Encarnacion for the first two and Steve Bero for the third.

grounds for criminalization become decisive. There are multiple normative pressures supporting greater punishment, like prevention and desert. Here I have been concerned with a countervailing pressure pulling punishments downward—the duty to not be cruel. But this normative pressure can only drag punishments down so far before the upward pressures overpower it. Wherever this equilibrium lies is where the legislature’s withholding of punishment to avoid being cruel should cease.

Finally, some worry that mercy conflicts with justice, and an analogous concern might afflict my argument.⁴⁶ The worry for mercy is this. There are reasons, sounding in retributive justice, to criminalize and punish culpable conduct. In failing to punish culpable conduct—even for kind-hearted reasons like mercy—we fail to satisfy the demands of retributive justice. We are not being fair to the people in the jurisdiction whose rights and interests the criminal law seeks to protect. Just as ratcheting down punishments out of mercy might seem unfair to those to be protected by the criminal law, would not the same be true for ratcheting down punishments to avoid being cruel? Thus, one might worry that this approach also conflicts with retributive justice.

In response, I accept that ratcheting down punishments to avoid being cruel does come at the expense of one kind of fairness. It departs a bit from the retributivist ideal of punishing to the extent warranted on culpability grounds. But it does so for reasons that are *internal to fairness*. This is an intra-fairness issue. My argument amounts to sacrificing a bit of fairness (captured in the ideal of retributive justice) due to the concerns of another closely related kind of fairness—what we might call *civic justice* (or a kind of equity), which reflects broader principles of political morality and good governance, and the breach of which also generates complaints and blame.⁴⁷ The view I have been articulating begins with retributive justice but then recognizes that additional fairness reasons pull the appropriate punishment levels down in places. This is a departure from one kind of fairness in order to satisfy another fairness concern—namely, the legislature’s duty to avoid the legitimate complaints it would face from legislating in cruel, callous, or vindictive ways.⁴⁸

46 For discussion of the conflict between mercy and justice, see Duff, who suggests that “mercy involves hindering the achievement of the goals that punishment serves” (“Mercy,” 474). See also Tasioulas, “Mercy.”

47 Cf. Duff, “Mercy,” 481–82. Duff suggests that mercy can sometimes function as “justice-completing equity,” which makes up for generally just criminal law rules that go awry in particular cases because of the rigidity of the rules, even though such rigidity itself may be needed to send a clear message.

48 Something similar may be discernable in excuses. Consider an act that (1) satisfies the offense elements and (2) has no justification. From 1 and 2, we can conclude that the law

4. CONCLUDING REMARKS

In this paper, I have developed a new sort of solution to the legal luck puzzle, which applies in different ways to reckless endangerment and attempts. We have seen how the legislature's duty to not be cruel requires finding meaningful ways to withhold full punishment, and we have seen why doing so when the wrongdoing causes no harm is a particularly good way to strike a balance between the competing normative pressures on the legislature. Withholding full punishment in cases of harmless wrongdoing is an especially good way to minimize the frustration of victims' claims to reaffirmation of rights while complying with the legislature's duty to not be cruel. My argument leaves open that there might be other institutional forms that can satisfy this legislative duty as well, but it at least provides an adequate normative justification for the legislature to ratchet down punishment for some categories of misconduct based on the degree of harm caused. The legislature may opt to deploy this technique for satisfying its duty to avoid cruelty in different ways for different forms of wrongdoing—such as by fully withholding punishment from harmless endangerments but imposing reduced criminal liability on intentional wrongs that luckily prove harmless (i.e., mere attempts). In this way, the legislature has a plausible normative rationale it can use to justify some luck in the criminal law.⁴⁹

University of Surrey, School of Law
a.sarch@surrey.ac.uk

deems the act to be all-things-considered wrong. But suppose that (3) the act is excused—meaning there are reasons of compassion or sympathy that give reason not to punish it fully. Perhaps a young person's father threatens her with severe beatings unless she kills someone. The killing is unjustified but may well be excused, at least partially. It arguably is a culpable wrong, but one where there are good fairness reasons *outside* of retributive justice—reasons of fairness to the wrongdoer stemming from compassion and sympathy for her plight—to punish less than the amount we would be entitled to impose on culpability grounds. This is an individual analog to the “don't be cruel” argument that I am suggesting applies at the societal level. Sacrificing a bit of retributive justice for other fairness reasons is a familiar move in the criminal law.

49 Many thanks to Steve Bero, Mihailis Diamantis, Hasan Djinder, James Edwards, Erik Encarnacion, Joe Horton, Ambrose Lee, Liat Levanon, James Manwaring, Andrew Simester, John Tasioulas, Chris Taggart, and audiences at Oxford University, King's College London, and the University of Iowa for extremely helpful feedback on earlier drafts of this paper.

REFERENCES

- Alexander, Larry, and Kimberly Ferzan. *Crime and Culpability*. Cambridge: Cambridge University Press, 2009.
- Alexander, Michelle. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: The New Press, 2010.
- American Law Institute. *Model Penal Code: Official Draft and Explanatory Notes*, May 24, 1962. Philadelphia: The Institute, 1985.
- Barrozo, Paulo. "Punishing Cruelly: Punishment, Cruelty, and Mercy," *Criminal Law and Philosophy* 2, no. 1 (January 2008): 67–84.
- Berman, Mitchell N. "The Justification of Punishment." In *The Routledge Companion to Philosophy of Law*, edited by Andrei Marmor, 141–56. New York: Routledge, 2012.
- Binder, Guyora. "Victims and the Significance of Causing Harm." *Pace Law Review* 28, no. 4 (Summer 2008): 713–37.
- Boeglin, Jack, and Zachary Shapiro. "A Theory of Differential Punishment." *Vanderbilt Law Review* 70, no. 5 (2017): 1499–559.
- Christopher, Russell. "Does Attempted Murder Deserve Greater Punishment Than Murder? Moral Luck and the Duty to Prevent Harm." *Notre Dame Journal of Law, Ethics and Public Policy* 18, no. 2 (2004): 419–35.
- Duff, R. A. "Mercy." In *The Oxford Handbook of Philosophy of Criminal Law*, edited by John Deigh and David Dolinko, 467–92. Oxford: Oxford University Press, 2011.
- Edwards, James, and Andrew Simester. "Crime, Blameworthiness, and Outcomes." *Oxford Journal of Legal Studies* 39, no. 1 (Spring 2018): 50–73.
- Husak, Doug. *Overcriminalization: The Limits of the Criminal Law*. Oxford: Oxford University Press, 2007.
- Jackmanm Tom. "Sessions Takes Federal Crime Policy Back to the '80s." *Washington Post*, May 12, 2017. <https://www.washingtonpost.com/news/true-crime/wp/2017/05/12/sessions-takes-federal-drug-policy-back-to-the-80s/>.
- Kaplan, John, Robert Weisberg, and Guyora Binder. *Criminal Law: Cases and Materials*. 7th ed. New York: Wolters Kluwer Law & Business, 2012.
- Kilpatrick, D. G., C. L. Best, L. J. Veronen, A. E. Amick, L. A. Villeponteaux, and G. A. Ruff. "Mental Health Correlates of Criminal Victimization: A Random Community Survey." *Journal of Consulting and Clinical Psychology* 53, no. 6 (December 1985): 866–73.
- Lackey, Jennifer. "Group Knowledge Attributions." In *Knowledge Ascriptions*, edited by Jessica Brown and Mikkel Gerken, 243–69. Oxford: Oxford University Press, 2012.

- LaFave, Wayne. *Substantive Criminal Law*. Vol. 2. West's Criminal Practice Series. 3rd ed. Eagan, MN: Thomson West, 2017.
- Lewis, David. "The Punishment That Leaves Something to Chance." *Philosophy and Public Affairs* 18, no. 1 (1989): 53–67.
- List, Christian, and Philip Pettit. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press, 2011.
- Morse, Stephen. "The Twilight of Welfare Criminology: A Reply to Judge Bazelon." *Southern California Law Review* 49, no. 6 (September 1976): 1247–68.
- Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.
- Sarch, Alexander. *Criminally Ignorant: Why the Law Pretends We Know What We Don't*. New York: Oxford University Press, 2019.
- . "Knowledge, Recklessness and the Connection Requirement between Mens Rea and Actus Reus." *Penn State Law Review* 120, no. 1 (2015): 1–51.
- Simester, A. P. "Is Strict Liability Always Wrong?" In *Appraising Strict Liability*, edited by A. P. Simester, 21–50. Oxford: Oxford University Press, 2005.
- Tadros, Victor. *Criminal Responsibility*. Oxford: Oxford University Press, 2005.
- Tasioulas, John. "Mercy." *Proceedings of the Aristotelian Society* 103, no. 1 (June 2003): 101–32.
- Westen, Peter. "An Attitudinal Theory of Excuse." *Law and Philosophy* 25, no. 3 (May 2006): 289–375.

EXPECTING EQUALITY

HOW PRENATAL SCREENING POLICY HARMS PEOPLE WITH DISABILITIES

Athmeya Jayaram

AT VARIOUS STAGES in the reproductive process, doctors offer prospective parents the opportunity to screen their embryos or fetuses (hereafter “embryos”) for genetic conditions that lead to future disability, such as cystic fibrosis and Down syndrome. When the screening comes back positive, the prospective parents—or “screeners”—then have limited information about the embryo: that it has a genetic condition and will have a future disability, or at least that it has a certain risk of developing one. If the screeners decide to terminate the embryo based on only this information, disability theorists argue, they are sending a harmful message to existing people with these conditions and disabilities. This is called the “expressivist objection” to screening and termination.

What message does it send? Defenders of the expressivist objection argue for two possibilities. S. D. Edwards offers an example of the first:

Consider a person currently living with cystic fibrosis. Such a person might hold the view that prenatal screening for cystic fibrosis, with a view to termination on grounds of the presence of cystic fibrosis in the fetus, sends a negative message to the person to the effect that it would have been better had he not been born.¹

In this example, terminating the embryo sends the message that, once it is known that an embryo has cystic fibrosis, it would be preferable that the embryo did not live than that it live with that condition.² This is supposed to imply that people with cystic fibrosis do not have lives worth living, which is a psychologically harmful message to people currently living with that condition.

1 Edwards, “Disability, Identity and the ‘Expressivist Objection,’” 418.

2 I use the neutral term “it” to avoid taking a position on whether embryos or fetuses are persons.

It can also be materially harmful if it strengthens the attitude that the lives of people with this and other conditions are not worth saving or extending, when the opportunity arises.

A second possible message seems initially less harmful, but may be more so because it is more relevant to daily life. Janet Malek, Adrienne Asch, and others argue that terminating an embryo based on a single characteristic sends the message that it is appropriate to evaluate and make significant decisions about a person (or future person) based solely on their disability. There are two ways in which one might arrive at this message. First, one might straightforwardly think that the disability is all one needs to know to evaluate the future person. As Malek puts it, “selecting against a future child on the basis of a disability signals that a disabling trait can be so significant and so undesirable that it eclipses all of the individual’s other traits.”³ Second, one might mistakenly think that a person who has one disability also has a host of others, which would lead one to conclude that a single disability is sufficient to evaluate a person. As Asch says, the “rehabilitation literature is full of examples of how able bodied people think of disabled people not as having specific disabilities, but as being *generally* incompetent.”⁴ In either case, however, the harmful message is supposed to be sent when an embryo is evaluated on the basis of the genetic condition alone. This message is harmful because it suggests that one should also evaluate the worth of current people with disabilities based solely on their disability, and not the many other aspects of their lives and character.

Theorists have offered three kinds of responses to the expressivist objection, each of which denies that screening and termination send any harmful message to people with disabilities. The first response argues that, as long as the screeners *could* be motivated to terminate the embryo by something other than the harmful judgments, the action does not send a harmful message. A second response claims that the action of termination is targeted at the genetic condition but not at the people who have it, so it does not imply anything about the worth of those people. A third response argues that the termination is motivated by the high costs of raising a child with a disability, rather than a judgment about people with that disability.⁵

3 Malek, “Deciding against Disability,” 217.

4 Amundson, “Disability, Handicap and the Environment,” 114n8, quoted in Asch, “Disability Equality and Prenatal Testing,” 329.

5 A fourth possible response concedes that screening and termination send a harmful message but that the harm is outweighed by its benefits. I will briefly address this concern at the end of the paper. However, this fourth response is largely in line with my argument, which is that the screening and termination of embryos with less severe genetic conditions sends a harmful message to people with those conditions, and that this harm gives us a *pro*

I will argue that the first two responses fail to answer the expressivist objection once we correctly understand when actions send harmful messages to their targets. It is not when the actor *could* have a harmless motivation or when the recipients identify with the targeted category. Instead, it is when a “reasonable person” would see the action as motivated by a harmful attitude.

However, the third response—that screening and termination may be motivated by cost—is reasonable. It is reasonable to see the screeners as motivated by the costs of raising a child with the disability rather than by a judgment about the disability or the people who have it. Nevertheless, this motivation still sends a message—different from the two suggested by disability theorists, but one that still causes harm. The message is that it is permissible not to pay a higher cost to support people with disabilities when there is a less costly abled alternative. This message echoes, sanctions, and reinforces the same attitude among public officials and employers, which has long motivated their refusal to provide people with disabilities with equal opportunity. In addition to affecting the rights and welfare of people with disabilities, this message also affects their sense of self-worth, or what John Rawls calls the “social bases of self-respect.”⁶

I further argue that this message is sent most clearly and harmfully by the state when it allows screeners to terminate because of cost considerations. It is reasonable to see this state action as motivated by the harmful attitude above for two reasons. First, as I will argue, the US government has a long history of failing to pay a higher cost to support people with disabilities, so it is reasonable to see the same motivation behind the refusal to regulate screening and termination. Second, there is no other reasonable and legitimate motivation for the government to allow screening and termination. As I argue through an analogy with sex-selective termination in India, it is not reasonable to see the government as motivated to respect reproductive rights because there is no right to information about one’s embryonic offspring. Nor, in certain cases, is it reasonable to see the government as merely serving citizens’ interests in avoiding the high cost of raising a child with a disability. Citizens only have this interest because of the government’s failure to fulfill its duty to provide equal opportunity to people with disabilities, so this failure cannot serve as a justification for further government action.

Now, I say “certain cases” because there are genetic conditions like cystic fibrosis that are so severe that no amount of government effort could equalize the costs to parents of raising a child with those conditions. (I will call such

tanto reason to stop screening for those conditions. I will therefore focus on evaluating the first three responses to the expressivist objection, all of which deny that screening and termination send a harmful message.

6 Rawls, *A Theory of Justice*, 386, and *Political Liberalism*, 318.

conditions “severe” in reference to the fact that the government cannot equalize the costs of raising a child with that condition.) In such cases, the government is not failing in its duty, so it is reasonable to see the government’s motivation as reducing the costs to prospective parents. For that reason, when the government allows the termination of embryos with severe genetic conditions, it does not send a harmful message to people with those conditions.

However, since the government *can* significantly reduce the costs of raising a child with a less severe disability such as Down syndrome, allowing screening and termination of these conditions does send a harmful message. (I will call such conditions “less severe” in reference to the fact that the government *can* (and should) equalize the costs of raising a child with that condition.) What message does the state send in allowing screening for less severe conditions? It sends the same harmful message that is evident in so many other government actions: Why pay the higher cost of supporting a person with disabilities, when you can just wait for an abled alternative?

1. FIRST RESPONSE TO THE EXPRESSIVIST OBJECTION: OTHER POSSIBLE MOTIVATIONS

The expressivist objection argues that prenatal screening and termination based on the genetic condition of an embryo sends a damaging message to people with that condition. To evaluate this claim, then, we need to know when an action sends a message. Allen Buchanan argues that an action (or decision) only sends a message when the action “presupposes” that message, either rationally or as a necessary element in one’s motivation.

Presumably, to say that a decision expresses (or presupposes) a judgment is to say either (a) that (as a matter of psychological fact) one could be motivated to make a decision of this sort only if one subscribed to the judgment (and that hence one couldn’t make the decision if one did not believe to be true what the judgment affirms), or (b) that one cannot rationally make the decision without believing what the judgment affirms.⁷

Under Buchanan’s view of what it is for an action to send a message, an action does not send a harmful message as long as we can think of a psychologically possible motivation, or rational alternative justification, for that action that does not send a harmful message. Buchanan then suggests several alternative motivations for terminating an embryo with a disability that do not send a

7 Buchanan, “Choosing Who Will Be Disabled,” 30.

harmful message: “one may simply wish to be spared avoidable and serious strains on one’s marriage or on one’s family. Or one may wish to avoid putting additional pressure on limited social resources to support disabled individuals. . . . [Or one may] desire not to bring into the world an individual with seriously limited opportunities.”⁸ Screeners who act on these beliefs or motives do not necessarily express a claim that certain lives are not worth living or that lives can be evaluated solely based on a genetic disability.

However, this is an overly demanding requirement for an action to send a message. There is always another way to interpret an action that is psychologically possible and rational, and avoids attributing a harmful motivation to the actor. For instance, rather than attribute a sincere motivation to the offending actor, one can always see the action as ironic, sarcastic, or a parody, which completely changes the motivation we attribute.

This is especially true of actions that send a harmful message to the disadvantaged because the action can often be interpreted as a *concern* for the disadvantaged, as in calls for assimilation, racial passing, or conversion therapy. Many such calls clearly send a damaging message to members of the targeted group: that the way they are is not acceptable. However, it is always *possible* to interpret the sender’s motivation as a concern for the targeted group. On this interpretation, the sender might be saying: “There is nothing wrong with being gay but, in our society, gay people face a life of discrimination so, out of concern for them, I think they should be “converted.” This may be a misguided way of showing concern for gay people, but it is not irrational to see the sender as motivated in this way. So, unless we concede that calls for conversion do not send a harmful message, we need a different test than whether there is a harmless motivation that is psychologically or rationally possible.

If the problem with Buchanan’s test is that it is always possible to reinterpret the motivations behind an action, then perhaps we should focus on the *actual* motivations of the sender. We could then say that if (and only if) the sender was actually motivated by a harmful judgment toward gay people then their actions send a harmful message. But this test goes too far in the other direction. A message can be sent regardless of the actual motivations of the actors involved. James Nelson offers the example of a group of people who raise the Confederate flag over the South Carolina State House. In this case, it does not seem to matter what their actual motivations are; the flag has a socially accepted meaning as a symbol of slavery and it is being used in a standard way to express that meaning.⁹ The flag raisers may actually be motivated by nostalgia or respect

8 Buchanan, “Choosing Who Will Be Disabled,” 31.

9 Nelson, “Prenatal Diagnosis, Personal Identity, and Disability,” 216.

for those who died in the Civil War. They may even, as I suggested earlier, be motivated to criticize the state's racist policies by associating the state with a racist symbol. But the very fact that this action *would* associate the state with racism means that the motivations of the actors do not matter to the message conveyed, since the message would be received as a racist one regardless of their actual motivations. The test of whether a message is sent must depend not on the sole possible or actual motivation for the action, but on how the motivation would be perceived.

But perceived by whom? How people interpret the display of the Confederate flag likely depends on whom you ask. A group of white Southerners may answer differently than a group of African Americans. This problem would be more pronounced in an even more racist time and place, like the antebellum South. At that time, the majority of people may have seen holding slaves as motivated by economic survival or ambition rather than racism. The perspective that interprets the message of an action should not simply reflect the social conscience of a society at the time. If it did, then a society in which the disadvantaged had internalized their lower status would not recognize any action as sending a harmful message. Instead, the perspective that attributes motivations to actions should be well-informed about how the proposed attitudes and actions fit into a larger historical pattern.

This more informed perspective will often be the perspective of the potential target of the message, since they are more likely to be aware of the pattern of actions and attitudes and less likely to try to rationalize away discriminatory attitudes. For example, as Sophia Moreau writes in trying to redefine the concept of "accommodations," there are many aspects of everyday life—such as the building of stairs instead of ramps—that are created to accommodate the abled in a way that may seem trivial to the designers and to most users.¹⁰ Just one less entrance accessible to people with disabilities, one fewer job. Given what they see as the small scale of the benefit, they may see themselves as motivated by a cost-benefit analysis rather than by any harmful indifference to the interests of people with disabilities. However, people with disabilities are much more likely to see the significance of the indifference behind these decisions because they experience a pattern of indifference, or worse, in so many other aspects of their lives.

There are two advantages to privileging the perspective of the targeted group. First, people with the targeted conditions will witness many more actions toward people like themselves by people like the senders and will therefore be in a better position to discern a pattern or common motivation behind

10 Moreau, "Discrimination and Subordination."

those actions. Where a screener may make only one decision concerning people with Down syndrome in a lifetime, a person with Down syndrome witnesses a lifetime of other people's decisions toward people like him.¹¹ Second, beyond personal experience, people with genetic conditions and disabilities are more likely to be aware of the history of treatment of people like themselves by people like the sender. People with disabilities are therefore better able to judge whether there is a historical pattern that provides evidence of the senders' likely motivation.

However, while this moves in the right direction, we should not take this too far. People with disabilities may have more experience with discrimination toward the disabled, but that volume of anecdotal evidence could also lead them to see discrimination too often; they may be understandably, but overly, sensitive. So, we need a standard that captures what is appealing about the perspectives of people with disabilities but provides some critical distance from that perspective as well.

Let us therefore call the desired perspective that of the "reasonable person." A reasonable person will not focus on the possible or actual motivations of the sender, but will think about how the sender's motivation would likely be seen. A reasonable person will further consider how the motivation would likely be seen in light of the historical pattern of actions and attitudes by people like the sender toward people like the target. This interpretation will be informed by, but not identical to, the targets' views of the action. A reasonable person will try to find the motivation that, in light of similar actions by similar people, as well as particular acts and statements by that person, best explains the action in question. In the case of genetic screening, we would want to know whether either of the harmful messages theorists have identified fits with the pattern of attitudes and actions exhibited by people like the screeners toward people with genetic conditions. If people like the screeners tended to show a great deal of concern for people with disabilities in their other actions, and even said so publicly, then it would not be reasonable to see them as motivated by a harmful attitude.

As always with "reasonable person" standards, privileging this perspective does not give us an obvious answer to whether screening and termination sends one of the harmful messages. To see whether it does, we will have to think more carefully about who is performing the action and what they (and people like them) have done in the past. However, identifying this perspective is an

11 Of course, it is also possible that people with conditions that impair cognition may not notice the patterns as well as others, which is another reason not to identify the right perspective too closely with the targeted group, but rather with a "reasonable person" who is aware of the experiences and history of the targeted group.

improvement on Buchanan's test. Buchanan denies that we ought to privilege any single interpretation of motive since there are many rational possibilities. If this were right, then even clearly harmful actions like advocating conversion therapy would not send any particular message. Privileging the perspective of the reasonable person helps us to identify clear cases of expressive harm, ones where any reasonable person would see an action as at least partly motivated by a harmful attitude.

The reasonable person standard also helps explain examples like the Confederate flag case. We can now say why it does not matter if the flag runners were actually anti-racist activists who sought to associate the state with a racist symbol. Unless it would be clear to a reasonable observer that the act was motivated by anti-racism—because, for instance, the flag raisers have an anti-racist track record and publicly stated their intentions—a reasonable person cannot be expected to see the act as it was intended. Instead, a reasonable person will see white people displaying the Confederate flag and come up with the most likely motivation for that action, based on a thorough knowledge of similar people performing similar acts. Based on all of this, a reasonable person would conclude that raising the Confederate flag appears to be motivated by racist attitudes, which therefore sends a harmful message to African Americans. By applying a similar analysis to screening and termination, I will argue that it does send a harmful message, but only one sent by the state to people with less severe disabilities. Before turning to that argument, however, I will consider two other responses to the expressivist objection.

2. SECOND RESPONSE:

TERMINATION TARGETS CONDITIONS, NOT PEOPLE

With this standard for when a message is sent, we can now apply it to the central question: Would a reasonable person see prenatal screening as sending a harmful message to people with the screened genetic conditions? Critics of the expressivist objection may still say no. For these critics, prenatal screening may imply a negative judgment on a genetic *condition*, but it need not say anything about the people who have that condition. Take cystic fibrosis (CF) for example. If you screen for CF and then terminate the pregnancy based on a positive result, you are clearly saying that you do not want a child with CF. But that does not express a judgment about people with CF any more than a flu vaccine expresses a judgment against flu sufferers. In the latter case, at least, a reasonable person would conclude that you are against the condition, not the people.

The first response to this objection is that there is a major difference between people with CF and flu sufferers; the flu is not identity constituting. Flu sufferers

do not think of themselves as such, and so would not take offense to a judgment against the flu. But at least some people with CF conceive of themselves as people with CF or, more generally, as people with a disability, so a judgment against the condition *is* also a judgment against the group of people who have it.

The problem with this response, the objector continues, is that it is not available to some defenders of the expressivist objection, such as Adrienne Asch. Asch argues that screening and termination send the second harmful message described above: that it is acceptable to evaluate people based on disability alone. The problem with that message, she argues, is that it reduces people to their disabilities. In making that argument, however, Asch seems to deny that disability is identity constituting. As Malek puts it:

Asch, the most consistent proponent of the expressivist argument, states that “disability is not, and need not, be either a ‘deep’ or a valued part of identity for everyone who shares the disability critique.” In fact, her primary objection to the use of [reproductive and genetic technologies] to prevent disability in future children is that such use suggests a reduction of disabled people to their disabilities. She therefore clearly *rejects* the idea that disabled individuals should be defined by their disabilities.¹²

If defenders of the expressivist objection must deny that disability is identity constituting, then they cannot use identity to differentiate between people with CF and flu sufferers. If this is right then, for both genetic conditions and the flu, targeting the condition does not target the person who has it.

This critique helps clarify the expressivist objection but is by no means fatal to it. Defenders can respond in several ways. The first is to affirm that disability is identity constituting but argue that this is consistent with the expressivist critique. This is most easily done by rejecting the second harmful message and focusing on the first. They may reject the suggestion that there is anything wrong with evaluating a person based on a disability and instead focus on the message that a life with a disability is not worth living. Since there is no contradiction between seeing *this* message as harmful and believing that disability is identity constituting, the expressivist objection still holds.

But one does not have to reject the second message to be consistent with the claim that disability is identity constituting. Defenders like Asch can insist that both messages are harmful, while holding that disability is only *partially* identity constituting. Recall that the second harmful message is that it is appropriate to evaluate a life based solely on a disability. Critics argue that rejecting this message is equivalent to rejecting the claim that disability is identity constituting.

¹² Malek, “Deciding against Disability,” 219.

There are, however, two ways in which disability might be identity constituting. In the first, a disability is the most important part of one's identity; it defines a person to the degree that it is appropriate to evaluate that person based on the disability alone. That is the sense of "identity constituting" that people like Asch must reject when they object to the second harmful message. But there is a second sense of identity constituting in which disability is *part* of one's identity but not necessarily the most important part. Unlike flu sufferers, people still see their disabilities as an essential part of who they are and would acknowledge that, without their disability, they would be a different person. But acknowledging all this is not equivalent to endorsing the message that it is appropriate to evaluate a life based solely on its disability; that would be overly reductive in the way the critics point out. Instead, defenders of the expressivist objection can maintain that disability is part of people's identities and so a part of understanding them, while objecting to evaluations based solely on disability.

So, one way to show that targeting a disability is not the same as targeting the flu is to claim that disability is (at least partially) identity constituting, and one can do so in a way that is consistent with the expressivist objection. But even if one denies that disability is identity constituting, there is another important difference between targeting the flu and targeting disabilities.

When people "target the flu" with a flu vaccine, they do not *see* the flu as identity constituting—as part of who people are. Contrast this with screeners, at least some of whom will see disability as a part of people's identity.¹³ So, when those screeners decide to terminate the embryo because of its future disability, they are more likely to see that disability as part of the future person's identity. In that sense, some screeners "target" the disability, in a way that flu vaccinators do not. So, even if we deny that disability is, in fact, partially identity constituting, if the screeners *think* that it is, then their action may still send a harmful message to people with disabilities. If screeners see disability as part of people's identities, a reasonable person would still recognize that they are motivated by a harmful attitude toward people with that disability.

To take another example, when parents terminate an embryo based on the fact that it would be their third child, no message is sent to third children everywhere because neither third children nor the parents see that characteristic as part of their identity. But in the case of disabilities, screeners may see disability as an essential aspect of those who have one and still decide to terminate based on that characteristic. So, if you share that characteristic, even if you do

13 For instance, significant numbers of Americans think people with Down syndrome should attend different schools or workplaces, suggesting that having Down syndrome is very important to how one is seen and treated; Pace, Shin, and Rasmussen, "Understanding Attitudes toward People with Down Syndrome."

not associate with others on that basis or consider them to be “like you,” you could object to the attitude expressed toward people with that characteristic. As Nelson argues, if we started screening out future bald men, the existing bald men may well complain, even if neither bald men nor others saw them as a group before.¹⁴ The strength of their complaint would depend on the degree to which others generally see bald men as a group, treat baldness as identity constituting, and tend to express negative judgments toward it—in other words, the degree to which a reasonable person would see the message as directed toward bald men. However, a message could be sent regardless of whether bald men see their baldness as part of who they are. This is the nature of identity: one cares about it whether it is self-conceived, socially ascribed, or some combination of the two. An insult to a socially ascribed identity is still an insult.

There are therefore two important ways in which, unlike the flu, targeting a disability also targets people with that disability. The disability may be part of people’s identity or, even if not, screeners may *think* that it is. Either way, the expressivist objection still holds.

3. THIRD RESPONSE:

TERMINATION TARGETS CONSEQUENCES, NOT PEOPLE

While the expressivist objection survives the first two responses, a third one is more successful. Critics may argue that, in screening and terminating, screeners are rarely motivated by a harmful judgment against particular disabilities or the people who have them. Instead, screeners are often motivated to avoid the *consequences* of a future child’s disability.

The expressivist objection claims that, when screeners find out about the embryo’s genetic condition, that condition is all that they know about the embryo. So, if they decide to terminate the embryo, they must be sending a message about the genetic condition and disability, since that is all they know about the embryo. However, critics may say, that is not strictly true. Screeners know about the presence of a genetic disability *and* the likely consequences of that disability. They know that it will require more time, money, and effort to raise that future child, and those are the considerations that motivate them—not the condition itself. If the screeners decide to terminate because of such ordinary concerns, critics may say, they are not sending a harmful message any more than any other parent who decides to terminate on similar grounds.

For example, if you consider the benefit of having a child to be worth spending a certain amount of time or money, and you come to find out that it will

14 Nelson, “Prenatal Diagnosis, Personal Identity, and Disability,” 219.

actually require a lot more, you may change your mind about having a child. As Nelson argues, the same may happen if you lose your partner or your job; your ability to provide for children may no longer be equal to the cost of raising them, and so you may change your mind.¹⁵ But, in doing so, you are not relying on a judgment about the worth of the potential child—just a judgment about the costs you are willing to bear. Many people do not want to have children at all, and they are not seen as sending a message to existing children.

What harmful message is sent by deciding that one does not want to pay a higher cost to bring a person with disabilities into the world? Whatever message it sends, it does not fit either of the messages that disability theorists have proposed. It does not imply that people with disabilities do not have lives worth living, since it is a judgment about whether to pay the cost of raising future children with disabilities, rather than a judgment on the quality of their lives. Nor does it seem to evaluate a life based on disability alone, since it focuses on the costs of that disability and how those costs might affect the parents' lives.

Nevertheless, there is a third kind of harmful message that theorists have not yet identified—one that is compatible with being motivated by cost. Consider an analogy with sex-selective termination. India and China (and many other developing countries) have had a long-standing problem with selective termination of female embryos and fetuses, which has caused an imbalance in the male-to-female ratio.¹⁶ This imbalance continues today, despite both countries passing laws making it illegal to find out the sex of the embryo before birth. There are multiple causes of sex-selective termination and they are hard to disentangle. The economic disadvantage of having a female child is both caused by, and reinforces, the cultural disadvantages that women suffer in these societies. Nevertheless, at least part of the cause is economic.¹⁷ Here are a few of the economic factors that screeners must consider:

Inheritance and land rights pass through male heirs, aging parents depend on support from men in the absence of national security schemes and greater male participation in the workforce allows them to contribute more to family income. Women, on the other hand, require dowries and leave the natal family upon marriage, which make them an unproductive investment.¹⁸

15 Nelson, "Prenatal Diagnosis, Personal Identity, and Disability," 218.

16 Abbamonte, "Sex-Selective Abortion in India"; and Gupta, "Return of the Missing Daughters."

17 Hesketh and Xing, "Abnormal Sex Ratios in Human Populations," 13272; Myers, "Sex Selective Abortion in India"; and Gupta, "Return of the Missing Daughters."

18 Barot, "A Problem-and-Solution Mismatch."

The lack of economic opportunities for women in these societies, along with the lack of social programs to address this inequality, make it far more costly for parents to have a daughter than a son.¹⁹

When screeners in these countries decide to terminate female embryos, it is therefore reasonable to think they are motivated partly by cost. But the judgment this motivation expresses is still a harmful one: it is permissible (and perhaps even preferable) not to bear the higher cost of supporting a female child when one could support a male for less. This attitude is both pervasive and harmful. Parents express this attitude when they invest in their sons' educations, while keeping their daughters out of school, often doing unpaid housework.²⁰ The result is the denial of equal opportunity in education and employment: only 68 percent of women can read or write, compared to 87 percent of men, and women make up only 25 percent of the labor force in India.²¹ Because of the social and economic conditions in these countries, women will not have equal opportunity unless the society agrees to bear the higher cost (because of the lower return) of investing in them. So, a judgment that women are not worth this higher cost when there is a male alternative is a judgment against equal opportunity.

Screening and termination based on disability sends a similarly harmful message, even when it is motivated by cost. Under current conditions, equal opportunity for people with disabilities will require society to bear a higher cost to support them, even when there is a less costly, abled alternative. For people with disabilities to receive an equal education and equal opportunity for jobs, school administrators and employers must have the opposite attitude: that they should bear the higher costs to provide "reasonable accommodation" even when they could support an abled person for less. This attitude is also important for the self-worth and rights of people with disabilities: it encourages them to disclose their disabilities to their employers without fear of being replaced, and then to sue for their rights if they are. The enforcement of the Americans with Disabilities Act, which requires employers to provide reasonable accommodation, depends on lawsuits to promote justice. If people are afraid to disclose their disabilities or get the message that society prefers not to pay a premium to accommodate them, they are unlikely to advocate for reasonable accommodation. This message also potentially implies that people

19 Gupta, Zhenghua, Bohua, et al., "Why Is Son Preference so Persistent in East and South Asia?"

20 Barcelos, "Child Gender and Parental Investments in India"; National Commission for Protection of Child Rights, *Vocational and Life Skills Training of Out-of-School Adolescent Girls in the Age Group 15–18 Years*.

21 Abbamonte, "Sex-Selective Abortion in India."

with disabilities are not entitled to other rights, such as access to health care. Equal access to health care requires employers or the government to pay a higher cost to promote the health and opportunity of people with disabilities, even when those same resources could provide for a greater number of abled people. In these ways, the unwillingness to pay a higher cost to support people with disabilities is almost as harmful to the rights, welfare, and self-worth of people with disabilities as a judgment that their lives are not worth living or that it is appropriate to evaluate a person based on disability alone.

The attitude that it is permissible not to bear the higher costs of supporting a person with disabilities is clearly harmful, but is it reasonable to attribute this attitude to the screeners? The attitude is certainly pervasive among decision makers in the United States. Employers, for instance, are often unwilling to pay the higher cost to accommodate people with disabilities, resulting in wide gaps in employment. Only 33 percent of people with disabilities of working age are employed versus 77 percent of people without disabilities.²² And the gap remains even among those who are clearly capable of cognitive work: 30 percent of college-educated people with disabilities are employed, versus 77 percent of college-educated people without a disability. The cause of this gap appears to be employers' attitudes toward applicants with disabilities, which is in turn caused by the requirement to provide them with reasonable accommodation.²³

Despite this widespread attitude, it may not be reasonable to attribute this harmful motivation to individual screeners. There is a more targeted motivation that seems equally reasonable but far less damaging. Rather than saying that individuals are motivated to screen and terminate by not wanting to pay a premium to support a person with disabilities, we may say that individuals are motivated by a narrower belief: that the lives of people with disabilities are not worth the extra cost when we have the option of *bringing into existence* a less costly individual. If this were the motive behind the action, then it would not send a harmful message to currently existing people with disabilities since it would not imply anything about whether *existing* lives are worth the extra cost. It would only be a judgment about how to evaluate future lives, one with a disability and one without.

The second issue with judging individual screeners is that, even if it were reasonable to attribute a harmful motivation to them, this only gives them a *pro tanto* reason not to terminate the embryo. They may still have a stronger reason

22 Bureau of Labor Statistics, "Persons with a Disability," 2021.

23 Ameri, Schur, Adya, et al., "The Disability Employment Puzzle"; and Jolls and Prescott, "Disaggregating Employment Protection."

that would make it ethically permissible, all things considered, to terminate, such as when the costs of raising a child with disabilities is prohibitively high. It may be a harmful message for an individual to say that the cost of supporting a person with a disability is too high, but when that cost is high enough, it may be too burdensome to expect individuals to refrain from sending that message. To return to the example of sex-selective termination, some poor families in India and China simply cannot afford to pay a premium to have a female child, so their reason to terminate the embryo may outweigh their reason not to send a harmful message. Similarly, it might be too burdensome to expect individuals to take on (what can be) a high cost of supporting a child with a disability. The Americans with Disabilities Act, for instance, only requires employers to provide “reasonable” accommodation when it is not an “undue hardship” on their business.

These are strong objections to assigning an ethical obligation to individuals to refrain from terminating embryos because of their disabilities. And they may be right; this is the inherent difficulty in determining when an ethical obligation is overly burdensome or in how a reasonable person would interpret a motive. Nevertheless, we get a much clearer answer when we consider another actor: the state.

4. THE STATE

States send a more harmful message to people with disabilities than individual screeners for two reasons: (1) it is more reasonable to see this message in the overall pattern of most states’ actions, and (2) unlike many individuals, states—at least the wealthier ones—have the means to avoid sending a harmful message. I will begin by arguing these claims with regard to sex-selective termination in India because most people will find it more intuitive that termination based on sex sends a harmful message. I will then argue that there is only one relevant difference between sex-selective termination in India and disability-selective termination in the United States. Because of this difference, I conclude, the US government only sends a harmful message when it allows the termination of *less severe* genetic conditions such as Down syndrome.

For both sex- and disability-selective termination, the argument will proceed as follows:

1. When the government allows selective termination, it performs an action that can send a message.
2. Based on the history of similar government actions, a reasonable person would see the motivation behind allowing selective termination as based on a harmful attitude.

3. There is no other plausible, legitimate, and harmless motivation that would be more reasonable to attribute to the government.

Beginning with the first claim, then: When the Indian government previously allowed sex-selective termination, was this an action? You might think it was not. After all, the government merely *allowed* doctors to provide information and patients to act on it, without endorsing either practice. This is similar to the government allowing free speech, without endorsing the content of that speech. Does the government send a message when it merely allows screening and termination? The answer depends on whether the government claims regulatory authority over the practice. The difference with free speech is that liberal governments have no regulatory power over (most) speech; it is beyond their purview. So, when the government does not regulate speech, it is not because it has evaluated the speech and allowed it to proceed. Rather, the government does not even consider the question of regulating it.

Initially, one may think something similar about screening and termination: just as the government respects the right to free speech by not considering whether to regulate it, the government respects the right to reproductive freedom by not considering whether to regulate it. However, the scope of free speech is far broader than reproductive freedom. Just about all speech is allowed because it is not the government's role to decide which speech should be free and which should be regulated. Reproductive freedom, on the other hand, has limits. Those limits are contested, but few think that the government has no regulatory role in deciding what kind of information prospective parents can have about a future child. If we could determine them, should we tell parents about the intelligence, athleticism, or beauty of a future child? One could argue over the wisdom of doing so, but the debate would not be about the parents' right to know.

In any case, both the Indian and American governments *do* claim regulatory control over selective termination—most obviously, in the Indian case, as the government has now prohibited parents from knowing the child's sex in advance.²⁴ If the government has regulatory control over the activity and chooses not to regulate it, it performs an action that can send a message.

What message it sends depends, as I have argued, on how a reasonable person would see the motive behind the regulatory action. And a reasonable person interprets motive based on what best explains this and other government actions toward women. A reasonable person would be aware that the Indian government has passed some laws to promote equal opportunity for

24 Andrews, Fullarton, Holtzman, et al., "Assessing Genetic Risks."

women, but has done very little to implement or enforce those laws.²⁵ A reasonable person would also be aware that there is much more that the government can easily do to promote equal opportunity in ways that would discourage sex-selective abortion: change inheritance practices, punish families for demanding dowries and dowry-related violence, incentivize families to have only daughters, as well as the basic requirement to enforce and monitor the success of the policies already on the books.²⁶ And, unlike individuals, the Indian government has the means to enact, fund, and enforce these policies on a broad scale. So, based on the history of inaction by the Indian government to promote equal opportunity for girls and women, a reasonable person would clearly attribute a harmful attitude to the state: that it is at least permissible not to pay the higher cost of providing equal opportunities to women when there are male alternatives.

How does this general pattern of actions and attitudes relate to sex-selective termination? Many prospective parents who decided to terminate based on the sex of the embryo did so because of this same harmful attitude: they were unwilling to pay the higher cost of raising a female child.²⁷ The Indian government had access to the same reports that this attitude was a major cause of sex-selective termination, but it neither did enough to reduce this financial motivation (by providing equal opportunity), nor did it enforce the ban on prenatal sex determination for many years.²⁸ What was the most reasonable explanation for these government (in)actions? It was not a concern for the reproductive freedom of its citizens. As I previously noted, the Indian government claims regulatory control over information like the sex of the embryo, so it did not consider selective termination a right that was beyond evaluation. Nor, unlike for many individuals, was the motivation the prohibitive cost of supporting female children. In this case, the cost to the government was merely the cost of prohibiting sex-selective termination, which was, and is, within its means. So, the most reasonable explanation for the government's permitting sex-selective termination is therefore the same explanation for its failure to

25 Basu, *Harmful Practices against Women in India*; and Menon-Sen and Kumar, *Women in India*.

26 Gupta, Zhenghua, Bohua, et al., "Why Is Son Preference so Persistent in East and South Asia?"; World Health Organization, *Preventing Gender-Biased Sex Selection*; and Abbamonte, "Sex-Selective Abortion in India."

27 See notes 20–21.

28 Abbamonte, "Sex-Selective Abortion in India," 31.

equalize opportunity for women: an attitude that it is permissible not to pay a higher cost to support equal opportunity for women.²⁹

Now, one might still object that there is another reasonable motivation behind the state action that does not send any message to women. One might argue that the government is not claiming that it is permissible to refuse to pay a premium to support women; it is simply furthering the interests of its citizens. Citizens have an interest in selective termination of female embryos because of the extra time, money, and effort required to raise them. They prefer not to pay the extra cost of bringing a female child into the world when there is a less costly male alternative. Because the state has a *pro tanto* reason to further the interests of its citizens, the state ought to make it permissible to screen and terminate based on sex. Since this motivation is purely to further the interests of citizens, it sends no harmful message to women about whether it is worthwhile to pay a premium to support them or whether the state endorses termination. However, this “neutral” justification is not available to the state, so it is not reasonable to see this as the state’s message. Consider, again, the steps of the justification:

1. In Indian society, women face unequal opportunities that make it more costly and difficult to raise a female child.
2. Because of these additional costs, some citizens have an interest in terminating female embryos.
3. The state ought not to interfere with actions that further the interests of its citizens, as long as failing to interfere does not conflict with any other moral requirement.
4. Therefore, the state ought not to interfere with the selective termination of female embryos.

The problem with this argument is that the societal inequality for women might be a good reason for individual citizens to perform (or refrain from) an action,

29 The state’s inaction may send an even more harmful message: that selective termination of embryos with disabilities is not just permissible, but desirable. Consider the analogous situation of employers. If an employer failed to provide equal opportunity to women by, let us say, not providing maternity leave, and the employer was aware that this led to fewer female employees, it would be reasonable to infer that the employer intended and welcomed this result. Similarly, when the state fails to provide equal opportunity for people with disabilities, it increases the cost to individuals of raising children with disabilities. It is reasonable to think that increased cost motivates individuals to terminate embryos with disabilities. If the state is aware that its inaction leads to selective termination, and continues to allow the practice, it is reasonable to think that the state intends this result. This sends the far more harmful message to people currently living with disabilities that their lives are unwelcome or discouraged by the state.

but it cannot be a good reason for the state to do so. This is because the state has a duty to redress that inequality. So, from the state's perspective, the fact of the inequality cannot serve as support for any actions other than redressing that wrong.

More generally, I am arguing that if a person or entity has a duty to right a wrong, and is in a position to do so, its failure to right the wrong cannot justify any further decision. This is because, if an entity has a duty, and is in a position, to right a wrong, and yet uses its failure to justify some further decision, the entity is acknowledging that it will not fulfill its duty. This acknowledgement is unjustified, so it cannot justify any further decision.³⁰

To make this clear, let me spell out the attempted argument:

1. The state has a duty, and is in a position, to address unequal opportunities for women.
2. The state is not going to perform its duty.
3. Because the state is not going to perform its duty, citizens have an interest in selectively terminating female embryos.
4. Because citizens have an interest in selectively terminating female embryos, and the state has reason to further its citizens' interests, the state has reason to allow selective termination.

Again, however, the state cannot make this argument because it cannot justify the second step: refusing to perform one's duty is not justified and so cannot transfer any justification to the conclusion. The only claim that a refusal can justify is an attempt to fulfill one's duty or to compensate those who are owed the duty.³¹

³⁰ This claim takes a position in the actualism versus possibilism debate. It is beyond the scope of this paper to defend my position in this debate, but I note it for those who hold the opposing view.

³¹ Here, one might object that there is a relevant difference between the state and an individual; the state is not a unified entity. So, while one part of the state fails to perform its duty to equalize opportunity, another part of the state may simply be responding to that failure. If this is right, then there is no internal contradiction when the Food and Drug Administration (FDA) allows screening and termination, because it is not acknowledging its failure to perform a duty, but is simply responding to the rest of the government's failure and trying to make the best of a bad situation. It is certainly possible that, while one part of a government is motivated by harmful attitudes toward women or people with disabilities, another part is clearly motivated only to help these groups. Perhaps the head of the FDA has a disability and is (shockingly) independent from the rest of the executive branch. In such cases, a reasonable person might see the FDA's motivation differently from the rest of the government's. However, as a consequence of the "reasonable person" analysis, such cases will be rare. Unless a reasonable person would have clear evidence to the contrary, we can attribute a harmful motivation to the FDA because it is reasonable to assume that FDA

So, government failure to equalize opportunity for women serves two purposes in this argument. It makes it reasonable to attribute a harmful message to the government in allowing sex-selective termination. And, it also makes it unreasonable to see the government acting on a harmless motive in which the government is simply furthering the interests of its citizens. Unless we can find an alternative justification for allowing sex-selective termination that is both reasonable and does not send a harmful message, there is a *pro tanto* reason not to allow the practice. It would send the harmful message to women that it is permissible not to pay a higher cost to provide them with an equal opportunity.

For the same reasons that the Indian government sends a harmful message to women when it allows sex-selective termination, the US government sends a harmful message to people with certain disabilities when it allows disability-selective termination. As before, the first step in the argument is to establish that when the US government allows disability-selective termination it acts in a way that can send a message. Like the Indian government, US government agencies like the Food and Drug Administration (FDA) and the Centers for Medicare and Medicaid Services claim regulatory authority over prenatal genetic testing, so leaving these decisions to doctors and patients is a *decision* not to intervene. More specifically, the FDA claims “enforcement discretion” to regulate genetic testing, which means it “has the authority to regulate tests but chooses not to.”³² And the government is clearly aware of the harm that choosing not to regulate genetic testing can cause: the National Council on Disability issued a report calling for more active regulation.³³ In some cases, the government approval is more explicit: Medicaid often covers the cost of genetic screening, which is a direct endorsement of its permissibility.³⁴ So, the US government’s refusal to regulate genetic screening and termination is an action that can send a message.

Next, I will argue that a reasonable person would see this government choice as sending a harmful message to people with certain disabilities for two reasons. First, there is a larger pattern of government action and inaction toward people with disabilities that suggests a common and harmful motivation: that it is permissible to fail to pay the higher cost of supporting people with disabilities when there is a less costly abled alternative. I have already discussed the lack of equal opportunity in employment, which is partly a result of poor government enforcement of laws requiring reasonable accommodation. One

officials are similar to other government officials, and the overall pattern of their actions is the same.

32 National Human Genome Research Institute, “Regulation of Genetic Tests.”

33 National Council on Disability, *Genetic Testing and the Rush to Perfection*.

34 Andrews, Fullarton, Holtzman, et al., “Assessing Genetic Risks.”

study found that “47 to 58 percent of accommodation-sensitive individuals lack accommodation and would benefit from some kind of employer accommodation to either sustain or commence work.”³⁵ And while progress has been made, many public buildings are still not accessible to people with disabilities.³⁶ For instance, two-thirds of schools still have physical barriers that limit access for people with disabilities, as do many forms of public transportation and public housing.³⁷ Most important for this discussion, there is also unequal (and likely inadequate) access to special education teachers in public schools, increasing the financial and emotional costs of raising a child with a disability, particularly for poor communities.³⁸

Second, there is no other legitimate and plausible motivation that would more reasonably explain the government’s allowing disability-selective termination. As before, the most obvious candidates are reproductive freedom and innocently serving the interests of citizens. Regarding reproductive freedom, the US government rejects any individual right to information regarding the genetic condition of the embryo when it claims the ability to regulate it. And the claim that the government is merely serving the financial interests of its citizens faces a similar problem as the analogous claim for sex-selective termination. For the government to claim this motive, it would have to acknowledge that it will not do its duty to equalize opportunity for people with disabilities, which would reduce the financial motivation for selective termination. Since the failure to perform a duty cannot justify any further action (other than compensation), it would not be reasonable to attribute this harmless motive to the government.

While the analogy generally holds for sex- and disability-selective termination, there are two potential differences. The first is that some may accept the claim that the government has a duty to equalize opportunity for women, but deny that it has a duty to equalize opportunity for people with disabilities. I cannot defend the latter claim here, but if the reader rejects it, then my analogy fails as well. The second difference is, however, a relevant one. As I have mentioned throughout, I am arguing only that the state sends a harmful message when it permits selective termination of embryos with *less severe* disabilities. I can now explain the reason for this limitation in the case of disability-selective termination.

35 Maestas, Mullen, and Rennane, “Unmet Need for Workplace Accommodation.”

36 Silvestrini, “The Americans with Disabilities Act at 30.”

37 US Government Accountability Office, “K-12 Education”; Piekarski, “Major American Cities Still Pose Problems for People with Disabilities.”

38 Mason-Williams, “Unequal Opportunities.”

The argument works for both sex- and disability-selective termination because the government's failure to equalize opportunities for women or the disabled makes the cost prohibitive for some screeners, leading them to terminate the embryo. This applies to all cases of sex-selective termination because it is *possible* for the government to equalize opportunities between men and women in all aspects of their lives. There are no inherent differences between the sexes that would explain the inevitable added costs of raising a daughter, so the government is theoretically able to equalize the costs and opportunities for both sexes.

However, this is not the case for people with severe genetic conditions. Some impairments are severe enough that, while the government can reduce the cost of raising a child with that impairment, it cannot make the cost non-prohibitive for many screeners. So, screeners who have embryos with severe genetic conditions will retain an interest in selective termination, whether or not the government fulfills its duty. In those cases, the government's justification does not rely on its own failure to perform its duty, which means there *is* a neutral justification available for allowing selective termination: serving its citizens' interest in avoiding prohibitive costs. So, when the government permits screening and termination of severe genetic conditions, it does not send a harmful message to people with those conditions, it merely recognizes that individuals have an interest in not bearing a prohibitive cost to raise a child—a cost the government cannot meaningfully reduce. The practical implications of this argument will depend on the details of the genetic condition in question and what the government can do to reduce the cost of raising a child with the resulting disability.

As an illustrative example, we can think of the difference between a fetal diagnosis of cystic fibrosis (CF) and one of Down syndrome (DS). The material and emotional cost of CF is enormous. A severe case is estimated to cost almost \$3,000 a month in health care costs and require two to three hours a day for treatment.³⁹ People with CF live an average of forty-seven years.⁴⁰ It would be impossible to equalize opportunity for people with and without CF, or to equalize costs to parents of children with and without CF, so individuals would have an interest in avoiding these costs regardless of the state's assistance. The government therefore has a harmless motivation for allowing screening and termination for CF.

Many cases of DS are quite different. The estimated out-of-pocket expenses average only \$84 per month, and many people with DS live fulfilling lives, with

39 Van Gool, Norman, Delatycki, et al., "Understanding the Costs of Care for Cystic Fibrosis"; Cystic-Fibrosis.com, "Managing Daily Routines and Treatment Schedules with Cystic Fibrosis."

40 Farber, Prieur, Roach, et al., "Difficult Conversations."

a life expectancy that is approaching the average American without a disability (sixty-five versus seventy-nine years).⁴¹ With a concerted government effort, one could imagine that life expectancy could eventually be similar. Even if not, it is not a significant added burden to individual screeners, since the child with DS is still likely to outlive the parents. The government could also narrow the additional cost in time and effort to raise a child with DS by providing educational and care assistance, in school or at home. If the government pushed to equalize opportunities for people with DS, screeners might no longer have an interest in avoiding the additional costs. In a case like this, there is no non-harmful motivation for allowing screening and termination. The only possible harmless motivation is invalidated by the government's failure to equalize opportunities for people with this disability.

Now, in making this argument, I have claimed that the state has a duty to provide equal opportunity to people with disabilities and is failing to fulfill that duty. One might therefore say that I have started with a big problem in order to point out a small one. The real problem is the state's failure to provide equal opportunity, so it is comparatively trivial that it sends a harmful message by allowing selective termination.

However, while it is certainly paramount for the state to fulfill its duties of justice, this conclusion still tells us something useful in the meantime. As long as the state fails to fulfill its duties, it has one less reason to allow actions that its failure incentivizes. It should not use its own failure to support people with disabilities as a reason to allow citizens to terminate embryos with disabilities. If and when the state fulfills its duties, however, then the state will no longer send a harmful message by allowing selective termination. This is an intuitive result: when a state displays more concern for people with disabilities, it is less reasonable to infer any harmful messages from its other policies.

Interestingly, however, when the state equalizes opportunities for people with less severe disabilities, the *individual screeners* may then send a harmful message to people with those disabilities. Because of state support, the screeners will no longer have a cost-based motivation to terminate embryos with less severe disabilities. If they continue to do so, then it may be reasonable to see their actions as sending one of the other harmful messages, such as the judgment that a life with a disability is not worth living. As the state displays more concern for people with disabilities, individuals will have to do so as well in order to avoid sending a harmful message.

41 Kageleiry, Samuelson, Duh, et al., "Out-of-Pocket Medical Costs and Third-Party Healthcare Costs for Children with Down Syndrome"; Rafii and Santoro, "Prevalence and Severity of Alzheimer Disease in Individuals with Down Syndrome"; Centers for Disease Control and Prevention, "Life Expectancy."

5. CONCLUSION

I have argued that there is a strong *pro tanto* reason for the government not to allow genetic screening and termination of less severe disabilities (where “less severe” is defined by whether the government can equalize the costs to parents of treatment). The *pro tanto* reason is this: allowing screening sends a harmful message to people with those disabilities that they are not worth the high cost when there is an abled alternative. I have proposed and rejected several considerations that would outweigh the expressivist objection, such as reproductive freedom and cost-based justifications. Nevertheless, there may be other reasons to allow screening and termination that would outweigh the damage of sending a harmful message. I cannot argue against further candidates here, but I will offer a final thought.

When countries like India and China restrict information on an embryo’s sex, they do so presumably because it would cause a sex-ratio imbalance that would detract from the quality of life of a certain number of citizens who would not find partners. As genetic screening becomes more sophisticated, these kinds of countervailing reasons will only become stronger. If information on traits like intelligence, athleticism, and beauty become available before birth, the societal cost of providing this information to parents increases. Among other concerns, we risk losing natural human variation and violating norms of distributive justice. And if distributional concerns can outweigh the value of full information, then it is plausible that concern about a harmful message that conflicts with equal opportunity can also outweigh the value of knowing about a future child’s (less severe) disability. As we think more about the value of genetic information, we may even come to see information as something that detracts from other values, such as those of unconditional love and acceptance.⁴²

Johns Hopkins University and University of Oxford
ajayara2@jhu.edu

42 I am grateful to the Institute for Practical Ethics at UC San Diego and to the Berman Institute of Bioethics at Johns Hopkins University for their support while I revised this article. Thank you also to audiences at the 2021 American Political Science Association Conference, the National Institutes of Health, and the Hastings Center for valuable feedback. Finally, I would like to thank Daniel Fogal and Adam Lerner for their comments on previous drafts. Adam, in particular, anticipated (and still holds) so many strong objections that the paper had to improve to keep up.

REFERENCES

- Abbamonte, Jonathan. "Sex-Selective Abortion in India: Estimates on the Occurrence of Sex-Selective Abortion in India and Some Possible Solutions to Eliminate the Practice." Population Research Institute, 2019. <https://www.pop.org/wp-content/uploads/2019/07/Sex-Selective-Abortion-in-India-Final-1.pdf>.
- Ameri, Mason, Lisa Schur, Meera Adya, Scott Bentley, Patrick McKay, and Douglas Kruse. "The Disability Employment Puzzle: A Field Experiment on Employer Hiring Behavior." National Bureau of Economic Research (September 2015). <http://www.nber.org/papers/w21560>.
- Amundson, Ron. "Disability, Handicap, and the Environment." *Journal of Social Philosophy* 23, no. 1: (March 1992): 105–19.
- Andrews, Lori B., Jane E. Fullarton, Neil A. Holtzman, and Arno G. Motulsky, eds. *Assessing Genetic Risks: Implications for Health and Social Policy*. Washington, DC: National Academies Press, 1994.
- Asch, Adrienne. "Disability Equality and Prenatal Testing: Contradictory or Compatible?" *Florida State University Law Review* 30, no. 2 (2003): 315–42.
- Barcelos, Silvio, Leandro Carvalho, and Adriana Lleras-Muney. "Child Gender and Parental Investments in India: Are Boys and Girls Treated Differently?" National Bureau of Economic Research (January 2012). <https://www.nber.org/papers/w17781.pdf>.
- Barot, Sneha. "A Problem-and-Solution Mismatch: Son Preference and Sex-Selective Abortion Bans." *Guttmacher Policy Review* 15, no. 2 (May 16, 2012). <https://www.guttmacher.org/gpr/2012/05/problem-and-solution-mismatch-son-preference-and-sex-selective-abortion-bans>.
- Basu, Asmita. *Harmful Practices against Women in India: An Examination of Selected Legislative Responses*. Addis Ababa, Ethiopia: United Nations Economic Commission for Africa, 2009.
- Buchanan, Allen. "Choosing Who Will Be Disabled: Genetic Intervention and the Morality of Inclusion." *Social Philosophy and Policy* 13, no. 2 (Summer 1996): 18–46.
- Bureau of Labor Statistics. "Persons with a Disability: Labor Force Characteristics 2020." February 24, 2021. <https://www.bls.gov/news.release/pdf/disabl.pdf>.
- Centers for Disease Control and Prevention. "Life Expectancy." April 9, 2021. <https://www.cdc.gov/nchs/fastats/life-expectancy.htm>.
- Cystic-Fibrosis.com. "Managing Daily Routines and Treatment Schedules with Cystic Fibrosis." September 2019. <https://cystic-fibrosis.com/daily-routine>.

- Edwards, S. D. "Disability, Identity and the 'Expressivist Objection.'" *Journal of Medical Ethics* 30, no. 4 (August 2004): 418–20.
- Farber, Julia Gray, Mary G. Prieur, Christine Roach, Rosemary Shay, Michelle Walter, Drucy Borowitz, and Elisabeth P. Dellon. "Difficult Conversations: Discussing Prognosis with Children with Cystic Fibrosis." *Pediatric Pulmonology* 53, no. 5 (2018): 592–98.
- Gupta, Monica Das. "Return of the Missing Daughters." *Scientific American* 317, no. 3 (September 2017): 80–85.
- Gupta, Monica Das, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung, and Bae Hwa-Ok. "Why Is Son Preference so Persistent in East and South Asia? A Cross-Country Study of China, India, and the Republic of Korea." *Journal of Development Studies* 40, no. 2 (2003): 153–87.
- Hesketh, Therese, and Zhu Wei Xing. "Abnormal Sex Ratios in Human Populations: Causes and Consequences." *Proceedings of the National Academy of Sciences* 103, no. 36 (September 5, 2006): 13271–75.
- Jolls, Christine, and J. J. Prescott. "Disaggregating Employment Protection: The Case of Disability Discrimination." National Bureau of Economic Research (September 2004). <http://www.nber.org/papers/w10740>.
- Kageleiry, Andrew, David Samuelson, Mei Sheng Duh, Patrick Lefebvre, John Campbell, and Brian G. Skotko. "Out-of-Pocket Medical Costs and Third-Party Healthcare Costs for Children with Down Syndrome." *American Journal of Medical Genetics Part A* 173, no. 3 (March 2017): 627–37.
- Maestas, Nicole, Kathleen J. Mullen, and Stephanie Rennane. "Unmet Need for Workplace Accommodation." *Journal of Policy Analysis and Management* 38, no. 4 (Fall 2019): 1004–27.
- Malek, Janet. "Deciding against Disability: Does the Use of Reproductive Genetic Technologies Express Disvalue for People with Disabilities?" *Journal of Medical Ethics* 36, no. 4 (April 2010): 217–21.
- Mason-Williams, Loretta. "Unequal Opportunities: A Profile of the Distribution of Special Education Teachers." *Exceptional Children* 81, no. 2 (January 2015): 247–62.
- Menon-Sen, Kalyani, and A. K. Shiva Kumar. *Women in India: How Free? How Equal?* United Nations Office of the Resident Coordinator in India, January 2001. <https://www.undp.org/india/publications/women-india-how-free-how-equal>.
- Moreau, Sophia. "Discrimination and Subordination." In *Oxford Studies in Political Philosophy*, vol. 5, edited by David Sobel, Peter Vallentyne, and Steven Wall, 117–46. Oxford: Oxford University Press, 2019.
- Myers, Christine. "Sex Selective Abortion in India." *Global Tides* 6, no. 3. (2012).

- National Commission for Protection of Child Rights. *Vocational and Life Skills Training of Out-of-School Adolescent Girls in the Age Group 15–18 Years*. National Colloquium Report, 2018.
- National Council on Disability. *Genetic Testing and the Rush to Perfection*. October 23, 2019. https://ncd.gov/sites/default/files/NCD_Genetic_Testing_Report_508.pdf.
- National Human Genome Research Institute. “Regulation of Genetic Tests.” September 25, 2020. <https://www.genome.gov/about-genomics/policy-issues/Regulation-of-Genetic-Tests>.
- Nelson, James Lindemann. “Prenatal Diagnosis, Personal Identity, and Disability.” *Kennedy Institute of Ethics Journal* 10, no. 3 (September 2000): 213–28.
- Pace, Jill E., Mikyong Shin, and Sonja A. Rasmussen. “Understanding Attitudes toward People with Down Syndrome.” *American Journal of Medical Genetics Part A* 152A, no. 9 (September 2010): 2185–92.
- Piekarski, Joseph. “Major American Cities Still Pose Problems for People with Disabilities.” Council for Disability Awareness, April 17, 2017. <https://blog.disabilitycanhappen.org/american-cities-problems-people-with-disabilities>.
- Rafii, Michael S., and Stephanie L. Santoro. “Prevalence and Severity of Alzheimer Disease in Individuals with Down Syndrome.” *JAMA Neurology* 76, no. 2 (February 2019): 142–43.
- Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1993.
- . *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press, 1971.
- Silvestrini, Elaine. “The Americans with Disabilities Act at 30: Progress, but More Needed.” *Legal Examiner*, July 24, 2020. <https://www.legalexaminer.news/2020/07/the-americans-with-disabilities-act-at-30-progress-but-more-needed>.
- US Government Accountability Office. “K–12 Education: School Districts Need Better Information to Help Improve Access for People with Disabilities.” GAO-20-448, June 30, 2020. <https://www.gao.gov/products/GAO-20-448>.
- Van Gool, Kees, Richard Norman, Martin B. Delatycki, Jane Hall, and John Massie. “Understanding the Costs of Care for Cystic Fibrosis: An Analysis by Age and Health State.” *Value in Health* 16, no. 2 (March–April 2013): 345–55.
- World Health Organization. *Preventing Gender-Biased Sex Selection: An Interagency Statement OHCHR, UNFPA, UNICEF, UN Women, and WHO*. 2011. <https://www.who.int/publications/i/item/9789241501460>.

CHILDREN, PARTIALITY, AND EQUALITY

David O'Brien

IT IS an unexceptional thought that parents must have latitude to be partial toward their own children—i.e., to act in a variety of ways that favor the interests of only their own children. The judgment that a just society accommodates some such parental partiality is apparently a fixed point of commonsense morality. But it is not obvious how this judgment is to be reconciled with a commitment to more general principles of justice, which seem to require impartial concern for others. I focus on this reconciliation problem as it arises for a liberal egalitarian theory of justice. Given its robust commitment to an ideal of equality, this theory faces special difficulties in accommodating the deliverances of commonsense morality concerning parental partiality. Nevertheless, the literature contains proposals that purport to effect such a reconciliation by *putting partiality first*—i.e., by subordinating a concern for equality to a concern for parental partiality. I criticize these proposals and suggest a different direction for reconciliation by *putting equality first*—i.e., by subordinating a concern for parental partiality to a concern for equality. This alternative reconciliation strategy, I argue, deserves to be taken seriously by liberal egalitarians. Whether it is the most plausible way to reconcile equality and parental partiality depends both on the relative moral weight of people's interests in parenting and in equal opportunity, and on how to measure the disvalue of unequal opportunity.

The paper is organized as follows. In section 1, I sketch the contours of the tension between equality and parental partiality. In section 2, I explain and criticize a prominent reconciliation proposal that puts partiality first. In section 3, I explain and criticize a reconciliation proposal from the recent literature that puts equality first in a weak sense, and I then motivate and defend a proposal that puts equality first in a stronger sense. In section 4, I respond to two objections concerning the practical implications of the equality-first reconciliation proposal defended in section 3. Section 5 concludes.

1. THE PARENTAL PARTIALITY PUZZLE FOR LIBERAL EGALITARIANS

Liberal egalitarianism, as I shall understand it, is distinguished by its commitment to a demanding equal-opportunity principle of justice, which condemns certain inequalities that do not appropriately reflect people's responsible choices—for example, inequalities that reflect only people's social background.¹ A puzzle arises when we try to apply this principle to societies that include both parents and children. The judgment that justice requires a demanding equal-opportunity principle of this kind to be satisfied is the first component of the puzzle.

The second component of the puzzle arises from the fact that well-off parents can benefit their own children in many ways that badly off parents cannot. Most obviously, they can give their children large quantities of income and wealth. Somewhat less obviously, they have the time and resources to benefit their children in a variety of ways—everything from expensive extracurricular lessons to reading bedtime stories—that play a significant role in improving their children's chances of having a good life. If well-off parents act in these ways, they seem thereby to cause the ideal of equal opportunity to be worse realized.²

The final component of the puzzle arises from the judgment that commonsense morality accords some latitude to parents to be partial. In a just society, it may seem that parents must have latitude to parent without constantly being required to look over their shoulder, so to speak, to check that their parenting is not violating a principle of justice. I deliberately use the neutral descriptor "latitude" here because, as discussed in sections 2 and 3, it is a substantive question how this commonsense judgment is most plausibly interpreted. But interpreting "latitude" straightforwardly, this judgment seems to imply permissions for parents—including well-off parents—to benefit their children in a wide variety of ways. Of course, it would not be plausible to believe that parents may do just anything for their children. Liberal egalitarianism is made *liberal* by according significant importance to people's rights—and so in a just society parents certainly may not, for example, inflict physical harm on others in order to benefit their own children. But commonsense morality seems to suggest that parents must be permitted to engage in at least some "protected" range of activities vis-à-vis their children. To take a favorite example from the recent literature, they must at least be permitted to read their children bedtime

1 See Temkin, *Inequality*, 12; Cohen, *Rescuing Justice and Equality*, 7, and *Why Not Socialism?* 17–22.

2 See Macleod, "Liberal Equality and the Affective Family," 223; Brighouse and Swift, "Legitimate Parental Partiality," 50; and Gheaus, "Personal Relationship Goods," sec. 2.3.

stories.³ And, importantly, part of the commonsense thought is that parents' permission to do so is not contingent on bedtime-story reading having only an insignificant distributive upshot. Indeed, the evidence suggests that such seemingly small-scale parental activities—activities that well-off parents have both more time and resources to engage in successfully—confer significant advantages on children.⁴

Putting these three components together yields the puzzle. To make the puzzle vivid, let us continue to take bedtime-story reading to stand in for the protected range of parental activities. If we are liberal egalitarians who wish to accommodate the deliverances of commonsense morality, we seem to believe both that (P₁) a society is just only if an equal-opportunity principle is realized, and (P₂) a society is just only if parents have latitude to read bedtime stories to their children. But then, as just noted, there is evidence that (P₃) if parents have latitude to read bedtime stories to their children, then the relevant equal-opportunity principle will not be realized. Put informally, the puzzle is that a robust commitment to equality, when taken seriously, looks like it threatens to reach into activities that seem beyond moral reproach—like parents reading bedtime stories to children—and give us reason to, say, tax, discourage, or (in the limit) prohibit such seemingly morally innocuous activities. The task of resolving this inconsistent triad P₁–P₃ is what I shall call the parental partiality puzzle for liberal egalitarianism.

I am going to set aside some ways of responding to the puzzle. One way out is to reject P₁ by rejecting substantive equality of opportunity entirely—e.g., by taking P₂ and P₃ to constitute a powerful argument against P₁.⁵ Another way out is to reject P₂ by rejecting parental partiality entirely—e.g., by taking P₁ and P₃ to constitute a powerful argument against P₂.⁶ Liberal egalitarians who wish to retain a commitment to the family in something like its familiar form, however, will want to investigate whether it is possible to *reconcile* equality and partiality.

3 Cf. Brighthouse, "Legitimate Parental Partiality," 57; and Segall, "If You're a Luck Egalitarian, How Come You Read Bedtime Stories to Your Children?" 24.

4 For an overview of the empirical evidence suggesting that parental activities like bedtime-story reading have a surprisingly large influence on children's lifetime welfare prospects, see Brighthouse and Swift, "Legitimate Parental Partiality," 58n23. Of particular note is Lareau's *Unequal Childhoods*, a study of how different, class-correlated parenting styles translate into differential advantages for children. (In section 4, I reconsider the important work of Lareau and others in connection with an objection to the reconciliation strategy that I propose.)

5 Compare an argument suggested by Nozick, *Anarchy, State, and Utopia*, 167–68.

6 Compare an argument discussed by Munoz-Dardé to the effect that taking a commitment to equal opportunity seriously might require the institution of the family to "change so significantly that we may not recognize it" ("Is the Family to be Abolished Then?" 55).

(To save words, in what follows I will sometimes say “partiality” instead of “parental partiality” and “equality” instead of “the relevant equal-opportunity principle or the state of affairs in which it is realized.”) I will be concerned with responses to the puzzle that propose such a reconciliation of equality and partiality.

Within this topic, I am also going to further restrict my focus. Some liberal egalitarians endorse a robust division of moral labor.⁷ On this kind of view, the scope of application of the equal-opportunity principle is restricted to how the major institutions of a society are arranged (i.e., to the so-called basic structure of a society); justice imposes no further requirements on individuals, including parents, other than to conform to and in various ways support the ways that principles of justice structure the institutions that are within the purview of the principles. If the scope of justice can plausibly be restricted in this way, this offers a straightforward basis on which to reconcile equality and partiality. On such a view, P_3 will turn out to be false simply because patterns of individual choice—including well-off parents choosing to benefit their children through activities like reading them bedtime stories—are not, in themselves, the sorts of things that fall within the scope of the relevant equal-opportunity principle. I am going to set aside this kind of response to the puzzle. I do so for two reasons. First, it is controversial whether such a robust division of moral labor can plausibly be sustained.⁸ Second, one might be motivated to accept a version of liberal egalitarianism that recognizes such a division of moral labor in part *because* one thinks that, otherwise, the demands of equality and partiality would unacceptably conflict.⁹ It is worth asking, first, whether the latter thought is correct.

7 Cf. Rawls, *A Theory of Justice*.

8 For the classic criticism, see Cohen, *Rescuing Justice and Equality*. Of those who dissent from Cohen's critique, some believe that the moral division of labor can be sustained, but in a way that accommodates some of the force of Cohen's critique by allowing that the demands of principles of justice are responsive, in some way, to patterns of individual choices in their economic and domestic lives (e.g., Neufeld, “Coercion, the Basic Structure, and the Family”; and Schouten, “Restricting Justice”). Others believe that the division of moral labor can be sustained in its strong form, where principles of justice are not responsive to patterns of individual choices in people's economic and domestic lives, given that individuals abide by their duties of justice to promote and sustain just institutions (e.g., Pogge, “On the Site of Distributive Justice”; Scheffler, *Equality and Tradition*; and Freeman, “The Basic Structure of Society as the Primary Subject of Justice”). It is only if the latter kind of response to Cohen succeeds that the puzzle could be straightforwardly dissolved by appeal to a division of moral labor.

9 Pogge, for example, writes: “Such a theory [that does not recognize a division of moral labor] has unwelcome implications. . . . The mandatory direct pursuit of [ends like equal

Before considering proposals for reconciliation, I should note another way that I have narrowed my focus. In framing the puzzle as I have done, I have in effect set aside some closely related puzzles. For example, Macleod discusses a similar tension between liberal equality and the family.¹⁰ But Macleod's puzzle is in two respects different from the puzzle that is my focus. First, Macleod's puzzle concerns not only how the family threatens to upset equal opportunity but how it threatens to upset children's claims as children—i.e., their claims with respect to their childhood.¹¹ Second, Macleod's puzzle concerns the conflict between the family and a particular version of resourcist egalitarianism—i.e., a version of egalitarianism that takes people to have claims to that which can be converted into welfare. In part because of these differences between our puzzles, Macleod can plausibly claim that a “reasonable but imperfect harmony” between resource equality and partiality can be achieved by designing institutions such that there is equality of basic resources among children.¹² I am focusing on a different puzzle both because I am interested in the prospects for reconciliation if resourcism about the currency of egalitarian justice is not assumed and because I want to ask whether “reasonable but imperfect harmony” marks the limit of how far reconciliation might go.¹³

2. PUTTING PARTIALITY FIRST

Suppose, then, that one wanted to reconcile equal opportunity and parental partiality. How might one do so? A natural thought is to propose that each of these values matters but that there is some principled basis for ordering them in cases in which they come into conflict. To derive such a basis, one might begin from the observation that, while it seems to be a fixed point in our judgments that parents must have latitude to read bedtime stories to their children, not just any parental partial activity seems to enjoy the same moral significance as this one. This observation suggests that if one could give an account of why parental

opportunity] may require citizens to violate any and all agent-relative goals (e.g., the goal that one's children flourish)” (“On the Site of Distributive Justice,” 161).

10 In Macleod, “Liberal Equality and the Affective Family.”

11 Cf. Macleod, “Liberal Equality and the Affective Family,” 222–23.

12 Macleod, “Liberal Equality and the Affective Family,” 230.

13 Other related puzzles include the special case of the tension between equal opportunity and parental school choice discussed in Swift, *How Not to Be a Hypocrite*; and Brighthouse and Swift, *Family Values*, ch. 5; for further discussion of this puzzle, see Macleod, “The Puzzle of Parental Partiality.” The tension between equality and parental partiality is also arguably a special case of a wider tension between justice and care and/or love; for discussion, see Gheaus, “How Much of What Matters Can We Redistribute?” and “Love and Justice”; and Cordelli, “Distributive Justice and the Problem of Friendship.”

partiality matters in the first place then this might yield a principled basis for distinguishing among parental activities, such that parents are permitted only to engage in those partial activities (saliently including bedtime-story reading) that are appropriately connected to the reasons why parental partiality matters. The thought is that one would thereby give partiality its due (since there would be a principled basis on which it *sometimes* takes precedence over equality) while still preserving a robust commitment to equality (since parental partiality would take precedence over equality *only* in that limited range of cases). Because the idea is to give parental partiality a certain precedence over equality, call this the reconciliation strategy that *puts partiality first*.

There are a variety of ways in which one might develop the strategy of putting partiality first. Each one corresponds to a different view about why parental partiality matters. For my purposes, it will be enough for now to focus on a paradigm instance of the strategy, the version developed by Brighouse and Swift.¹⁴ Focusing on that version of the partiality-first strategy will help to illustrate both the strategy's attractions and its limitations. At the end of this section, I return to the question of how my discussion generalizes to other versions of the strategy.

Brighouse and Swift's core idea is to delineate a special class of partial activities that have significant moral value and to claim that, while parents have latitude to engage in partial activities within this class, partial activities outside this special class may be taxed, discouraged, or (in the limit) prohibited. To make this cut among partial activities, Brighouse and Swift appeal to what they call family relationship goods (FRGs). FRGs are a subset of the ways that being in family relationships benefit parents and children—e.g., by satisfying children's interests in being loved and being able to develop as adults who can flourish, and by satisfying parents' interests in playing the special social role of being responsible for a child developing into an adult who can flourish.¹⁵ These are weighty goods—intrinsically and instrumentally valuable both for those who receive them and for third parties. And some things that parents do are, in ordinary circumstances, necessary to produce FRGs; indeed, when this is so, the parental activity will typically partly *constitute* an FRG.¹⁶ Reading

14 Brighouse and Swift, "Legitimate Parental Partiality" and *Family Values*. For a related proposal, see Macleod, "Liberal Equality and the Affective Family."

15 See Brighouse and Swift, "Legitimate Parental Partiality," 53–54.

16 In general, activities necessary to produce a good need not also partly constitute the good. So there is an interesting sub-question about whether the morally significant line is between activities that are *necessary* for producing a good and those that are not, or between activities that partly *constitute* a good and those that do not. Brighouse and Swift, as I explain in the text following this note, endorse the former view. But since the case of

bedtime stories to children is, again, a representative example. Were parents not to engage in this kind of activity, children could not develop as loved, well-rounded adults, and parents could not play the special role of an adult responsible for the child so developing; moreover, engaging in this activity is not just necessary for realizing an FRG but partly constitutes one. But not all parental activities are, in ordinary circumstances, necessary to produce FRGs. Leaving the entirety of one's extensive estate to one's children is, ordinarily, not necessary to produce FRGs if indeed it produces any. Brighouse and Swift thus propose that whether a parent's partial activity falls in the protected class depends on whether it is, in ordinary circumstances, necessary to produce FRGs. If and only if it is, then parents have latitude to engage in this kind of activity. The core of the Brighouse and Swift proposal, in other words, is the claim that there is a limit to the degree to which partiality, properly understood, licenses parental activities that disrupt equality. The hope is that one can thereby recover a robust commitment to equality (since many partial activities that would disrupt it are ones that parents will not be licensed to engage in) while retaining a robust commitment to partiality (since latitude is allowed for a variety of partial activities on the part of parents—in particular those, like bedtime-story reading, that commonsense morality recognizes as nonnegotiable).

The attraction of the strategy, then, is its promise of finding a well-motivated basis for blocking the way in which demands of equality apparently overreach, threatening apparently morally innocuous activities. That overreach is blocked by finding an important value that those activities serve. We can next ask how to classify the partiality-first strategy: Which of the inconsistent triad P_1 – P_3 does it reject? As Brighouse and Swift note, it would be unduly optimistic to think that taxing parental gifts and bequests (and other such parental activities outside the protected class) could entirely offset the effects of bedtime-story reading (and other such activities within the protected class).¹⁷ So the partiality-first strategy cannot be a rejection of P_3 on the grounds that there is no conflict between equality and what Brighouse and Swift call legitimate parental partiality. Even if one accepts Brighouse and Swift's cut among partial activities, some of those that are within the protected class, like bedtime-story reading, will have a substantial dis-equalizing effect on the distribution of opportunity.

bedtime-story reading (which constitutes, and is not merely a necessary condition for, an FRG) motivates their view, one might question whether their cut among partial parental activities is faithful to the motivating case. However, I set this line of objection to the view aside here.

17 Thus Brighouse and Swift summarize their view as the claim that “the family and equality do not conflict *nearly as much as* is commonly thought” (“Legitimate Parental Partiality,” 80, emphasis added).

In fact, there are at least three distinct versions of the partiality-first strategy to be distinguished. We can distinguish among them by asking the adherent of the strategy what attitude they would take were equal opportunity to be promoted in a way that encroaches on the protected class of parental activities. One possible response would be that there is *nothing* good about this—no reason of justice, or any other normative reason, that favors it. Another possible response would be that while there is *something* good about this, it would not be favored by a reason of justice. A third possible response is that while there is *a* reason of justice favoring this, it is outweighed by other reasons of justice. These are all versions of the partiality-first strategy, but they correspond to different views about how equal opportunity matters. The first response suggests a view on which the relevant kind of equal opportunity is a justice-related consideration but only a conditional one. The second response suggests a view on which the relevant kind of equal opportunity, although it is an unconditional consideration, is not (or not always) a justice-related one. The third response suggests a view on which the relevant kind of equal opportunity is a justice-related consideration that matters unconditionally but that can be outweighed by other justice-related considerations.¹⁸ Thus we have three distinct versions of the partiality-first strategy. The first version corresponds to rejecting P₃. The claim would be that the relevant kind of equal opportunity can be fully realized even when parents have latitude—in the sense of full-blooded permission—to read bedtime stories to their children. The second and third versions correspond to rejecting P₁. The claim would be that, while there is a genuine conflict between parents having these permissions and the relevant kind of equal opportunity, justice does not require that kind of equal opportunity. According to the second version, that is because the relevant kind of equal opportunity is not a justice-related consideration. According to the third version, that is because the relevant kind of equal opportunity, although a justice-related consideration, is not the only such consideration.

Each of these versions of the partiality-first strategy in some way curtails the importance of equal opportunity; that, of course, is the guiding idea behind the partiality-first strategy. But none of them is a wholesale rejection of equal opportunity. So each can still plausibly be claimed to count as a reconciliation strategy. They differ in how deep, as it were, they take the conflict between equality and partiality to go. Rejecting P₃ is the mark of taking the view that the two values, properly understood, do not conflict at all. Rejecting P₁ is the

18 There are in turn two sub-varieties of this third response: on one, the relevant kind of equal opportunity is lexically outweighed when it conflicts with partiality in this way; on the other, the relevant kind of equal opportunity is outweighed but non-lexically. This difference will not be important in what follows.

mark of taking the view that the two values, while they conflict, can be ordered with respect to justice in some way.

Granted that these are ways of reconciling equality with partiality, we can next ask whether the partiality-first strategy is a plausible way of solving the parental partiality puzzle for liberal egalitarianism. The discussion above, I believe, suggests grounds for thinking that liberal egalitarians have some reason to try to find a different reconciliation strategy. The second version of the partiality-first strategy requires liberal egalitarians to take (at least some) equalizing of opportunity to be a non-justice-related value. But this sits ill with the motivations for liberal egalitarianism.¹⁹ The first and third versions of the strategy entail that FRGs (and the parental partiality that these goods in turn license) are weighty enough either to vacate entirely the value of equalizing opportunity, to lexically outweigh it, or to significantly but non-lexically outweigh it. Although not self-contradictory or inconsistent, these are strong claims. And some liberal egalitarians may therefore wish to ask whether the price of reconciling equality and partiality must be the downgrading of equality's significance in one of these ways.²⁰ So putting partiality first is not obviously the best-motivated way for the liberal egalitarian to solve the parental partiality puzzle. There is some motivation, then, for the liberal egalitarian to ask whether there is an alternative reconciliation strategy that is both plausible and does not require equality to be downgraded in significance in this way.²¹ It is worth asking, in

19 For a helpful discussion of egalitarianism's animating thesis that there is a conceptual connection between justice and equality, see Arneson, "Justice Is Not Equality," 73–74; for similar claims, see Temkin, "Justice, Equality, Fairness, Desert, Rights, Free Will, Responsibility, and Luck."

20 Accepting the third version of the strategy, Brighouse and Swift write:

We have now offered . . . reasons not to pursue fair equality of opportunity all the way. . . . We must be prepared for children of similar talent and ability raised by different parents to enjoy somewhat *unfairly* unequal prospects of achieving the rewards attached to different jobs, *since the alternative would cost too much in terms of [FRGs].* (*Family Values*, 44, emphasis added)

I leave aside here the large task of *directly* assessing whether Brighouse and Swift's complex case for their strategy would justify these strong claims about FRGs' moral significance. The alternative reconciliation strategy that I discuss in section 3.2 is, in effect, an *indirect* challenge to Brighouse and Swift's case for their strategy. Brighouse and Swift also note, but dissociate themselves from, the first version of the strategy, which in effect "allow[s] conflicts to shape the very way that we understand the conflicting [values] themselves" (*Family Values*, 44).

21 It is worth noting that Brighouse and Swift also take FRGs to be themselves *among* the *distribuenda* of egalitarian justice. So their strategy might be claimed to still give equality a kind of priority in a different sense. But granted that their strategy allows a top-ranked

other words, whether it is possible to reconcile equality and partiality in a way that puts equality first.

Before continuing on, it is worth asking to what degree the considerations above motivate looking for an alternative to the partiality-first strategy in general, as opposed to Brighouse and Swift's particular version of it. The distinguishing feature of the Brighouse and Swift version of the strategy is the appeal to FRGS as the grounds of parental partiality. In particular, as noted above, theirs is a view on which FRGS, in turn, matter because they are significant constituents of both children's and parents' welfare. The literature contains a variety of alternatives to this view.²² The value of parental partiality might be grounded in something other than the relationship goods such partiality produces.²³ And the value of the relationship goods that partiality produces might be grounded in something other than their contribution to people's welfare.²⁴ In general, then, these views will differ from Brighouse and Swift's version of the partiality-first strategy both in where they locate the cut between the protected class of partial activities and in the grounds they proffer for locating the cut where they do. Consider, for example, a view on which the value of FRGS is exhausted by their contribution to giving people the capacities required to be autonomous.²⁵ If one adds a view on which the capacities required to be autonomous are modest, then such a view might license a much smaller protected class of parental activities than Brighouse and Swift's. Conversely, other views about the value of partiality—e.g., views on which children have moral rights to be cared for and on which the demands of care are extensive—might license a much larger protected class of parental activities.

For each of these views about why partiality matters, then, it would take significant further work to evaluate whether the corresponding version of the partiality-first strategy offers a plausible and principled basis for the liberal egalitarian to solve the parental-partiality puzzle. So, officially, the case for exploring an alternative to the partiality-first strategy that I have offered is a

outcome in which both equality and partiality are fully realized, in cases of conflict between these values, partiality is put first.

22 For a helpful overview of the landscape, see Gheaus, "Personal Relationship Goods."

23 For example, it might be grounded in the reasons given by associative duties (cf. Gheaus, "Personal Relationship Goods," section 2.3 and the references therein).

24 As Gheaus notes, the significance of relationship goods has been grounded in considerations as diverse as needs, autonomy, rights, the imperatives of care, and political citizenship ("Personal Relationship Goods," secs. 2.1, 3.2–3.6; see also Macleod, "How Not to Be a Hypocrite," 314).

25 On the connection between relationship goods and autonomy, see, e.g., Brownlee, "A Human Right against Social Deprivation."

merely presumptive one. However, there is, I believe, some reason to think that the considerations advanced above will apply to at least many of these alternative views.²⁶ Albeit on different bases, each of them purports to curtail the importance of equal opportunity. And so, as noted above, given the centrality of equal opportunity in the liberal egalitarian view of justice, there will be some cost to the view if equality and partiality can only be reconciled by putting partiality first.²⁷

3. PUTTING EQUALITY FIRST

I have suggested that liberal egalitarianism faces a puzzle in reconciling equality and partiality. And I have explored one attractive strategy for doing so, which effects reconciliation by putting partiality first. Its core claim is that, from the point of view of justice, certain kinds of partiality matter more than equality. But downgrading equality's importance in this way, I suggested, is in some tension with liberal egalitarianism in a way that motivates exploring whether the parental partiality puzzle might be solved by instead finding some basis for putting equality first. Now, just as the *prima facie* trouble about putting partiality first was the threat of coming unmoored from a robust commitment to equality, the *prima facie* trouble about putting equality first is the threat of coming unmoored from the latitude that commonsense morality affords to parents. I begin in section 3.1 by critically evaluating a strategy that seeks to avoid

26 An interesting class of exceptions are views on which the moral value of relationship goods consists, either wholly or partly, *in* their role in allowing people to enjoy equal opportunity. See, for example, Brownlee, "A Human Right against Social Deprivation"; Cordelli, "Distributive Justice and the Problem of Friendship"; Brake, "Fair Care"; and, for discussion, see Gheaus, "Personal Relationship Goods," sec. 3.7. For a related "moralized" view of parental partiality on which partiality is moral reason giving only if it does not violate background principles of justice, see Bou-Habib, "The Moralized View of Parental Partiality." If this is why relationship goods (and so, derivatively, partiality) matter, this suggests that partiality could not take precedence over equal opportunity when they conflict. But it is unclear what such views would imply about well-off parents engaging in (equal-opportunity-disrupting) bedtime-story reading. Since on such views partiality would matter only insofar as it bears on equal opportunity, the answer would seem to turn on how to aggregate the gains and losses of different increases and diminutions in the opportunity sets of children of different social backgrounds. If such gains and losses are to be aggregated in a straightforward additive way, such views threaten to fail to count as reconciliation proposals since they seem to be flat rejections of ν_2 . The equality-first strategy that I defend in sec. 3.2 offers, I believe, a more plausible basis for putting equality first.

27 I thank an anonymous referee for the journal for suggesting that I consider how my arguments would extend to other views about the value of partiality and the grounds of the value of relationship goods.

that threat by putting equality first in a weaker sense, and then—motivated by the discussion of this weaker equality-first strategy and the partiality-first strategy in section 2—I proceed in section 3.2 to propose a strategy that puts equality first in a stronger sense.

3.1. Putting Equality First in a Weaker Sense

Shlomi Segall has proposed an ingenious alternative strategy to Brighthouse and Swift's strategy for reconciling equality and partiality, which promises to accommodate the moral significance of FRGs without compromising equality's importance.²⁸ The core idea of Segall's proposal is to note that accepting the moral significance of FRGs does not commit one to counting partial activities like bedtime-story reading as permitted at the bar of justice. (When not otherwise indicated, "permitted" and "not permitted" in this subsection are to be understood as indexed to the subset of moral reasons that concern justice; to save words, I sometimes omit this qualifier.) To explain how it could then be that parents could nevertheless have *latitude* to do what is not *permitted*, Segall appeals to a distinction between justification and excuse. Actions that are permitted are *justified*, in the sense of not being opposed by the balance of (justice-related) reasons. But some other actions, while not permitted (because not in this sense justified), are nevertheless *excusable*. The mark of an action having the status of being excused is that, while the action is not permissible, performing it does not make one liable to blame or culpability of the kind that ordinarily attaches to performing impermissible actions.²⁹ This, Segall claims, is what can be true of well-off parents who read bedtime stories to their children. This strategy thus reconciles equality and partiality by rejecting P3. But it does so on a different basis than the partiality-first strategy. On Segall's view, no activity that compromises equality is permitted (at the bar of justice). Bedtime story reading does so. So it is not permitted (at the bar of justice). But it does not follow that parents do not have *latitude* to engage in it. They do: although not permissible, it is nonetheless the kind of activity that can be excusable. That, Segall adds, is enough to vindicate the judgment of commonsense morality that parents must have *latitude* to engage in such activities. To vindicate this judgment, he suggests, it would be enough to claim that, although these parents do what is impermissible (at the bar of justice), they are not culpable or blameworthy for doing so. So far as the permissibility of parents' actions go,

28 In Segall, "If You're a Luck Egalitarian, How Come You Read Bedtime Stories to Your Children?"

29 I take this helpful gloss on the distinction from Enoch, "The Masses and the Elites," 5–6. As Enoch notes, although the distinction is more familiar within the legal domain, it seems plausible that a similar distinction is also available within the moral and political domains.

then, this strategy puts equality first. But since the impermissibility of what they do does not come along with its usual consequences of blameworthiness and culpability, this strategy puts equality first in only a weaker sense. Call it the strategy that *puts equality weakly first*.³⁰

Segall's case for this strategy is straightforward. Segall assumes that, relative to the ideal of justice, equal opportunity matters more than partiality does. In reading bedtime stories to their children, well-off parents are compromising equal opportunity. So what these parents are doing could not be supported by the balance of justice-related reasons. Indeed, as Segall adds, given these assumptions, a view on which well-off parents were permitted *simpliciter*, at the bar of justice, to read bedtime stories to their children would appear to give parents guidance that is both morally and psychologically incoherent. That would be so not only because the balance of justice-related reasons does not favor engaging in bedtime-story reading, but because well-off parents would apparently then be required—when not under the special protection of the permission of partiality—to spend time and resources offsetting the effects of their bedtime-story reading, even though that activity is, by hypothesis, permitted.

We thus seem to have two plausible strategies for resolving the puzzle: one corresponding to the view that equality matters more than partiality, the other to the view that partiality matters more than equality. But putting equality weakly first is not, I shall next suggest, a successful reconciliation strategy. I have two objections to this strategy. First, I do not think that an adequate explanation is available for the claim that well-off parents could be excused, at the bar of justice, for their bedtime-story reading. In standard cases, an act is excused when it does not, in some morally significant sense, reflect the agent's stance or what she took herself to be doing. But well-off parents who are also well-informed parents know the effects of their bedtime-story reading. So they cannot be claimed not to know what they are doing—as, e.g., someone who fires what she mistakenly believes to be only a toy pistol does not know what she is doing. Segall suggests that, by appealing instead to these parents' intentions, we can explain how their acts can be excused.³¹ But, given that these parents know the effects of what they do, it is questionable whether this is sufficient to explain their having an excuse. Compare: an oil merchant who sells poisoned cooking

30 Segall writes that “the fact that [bedtime-story reading] is pursued for its own sake ... does *not* make it a just [activity and] does not even suffice to make bedtime reading an all-things-considered justified undertaking. My claim, rather, is that having the right motivation may sometimes *excuse* an otherwise unjust activity” (“If You're a Luck Egalitarian, How Come You Read Bedtime Stories to Your Children?” 34).

31 Segall, “If You're a Luck Egalitarian, How Come You Read Bedtime Stories to Your Children?” 34.

oil, intending to turn a profit, is no less culpable than a coffin merchant who sells poisoned cooking oil with the same intention, even though only the latter intends harm as a means to their end.³² The fact that one intends to produce a good does not, in other words, generate an excuse if the good that one aims at is significantly lesser in importance than a non-intended, but foreseen, harm. According to the equality-weakly-first strategy, partiality is much less important than equality relative to the ideal of justice. So the good that well-off parents aim at is significantly less important, at the bar of justice, than the bad that their acts cause. So what they do could not be excused at the bar of justice.

Segall might respond that it would be enough, to generate an excuse at the bar of justice, that someone aims at a non-justice-related good that is sufficiently important, relative to the standards of that wider or different domain. And it may be true that, *all things considered*—i.e., taking into account justice-related and moral-but-non-justice-related reasons—this kind of partiality is more important than equality.³³ But the claim that considerations from one domain can affect the evaluative status of an action, relative to a different domain, seems dubious. No matter the magnitude, relative to the aesthetic domain, of the *aesthetic* good at which one aims, one could not have a *moral* excuse for failing to save a drowning child in virtue of the fact that one was aiming at this aesthetic good when one failed to save the child's life.³⁴

32 For this illuminatingly vivid case, see Foot, "The Problem of Abortion and the Doctrine of Double Effect," 3.

33 That would be enough to make what parents do all-things-considered *permissible*. And this appears to be Segall's considered verdict about what well-off parents do: the claim is that when well-off parents do what is necessary to produce FRGs, that they do what is *morally* permitted—i.e., permitted relative to the domain that includes both justice-related and moral but non-justice-related reasons. Segall's full view, in other words, is that although partiality is not valuable enough to make reading bedtime stories *just*, it is valuable enough to make reading bedtime stories (unjust but) *permitted*, all things considered. (And, moreover, when parents are aiming *at* that valuable thing—and not at, what they are also doing, unjustly advantaging their children—then their activities are excused, *at* the bar of justice.) This would resemble one version of the partiality-first strategy discussed above. So Segall's view differs from this version of the partiality-first strategy only insofar as the latter strategy allows that partiality is valuable enough to make reading bedtime stories *just*. Rather than subordinating equality to partiality *within* the ideal of justice, in other words, Segall shifts the point of subordination elsewhere, subordinating *justice* (understood to include, centrally, equality) to *partiality*.

34 A quite different explanation for why parents are not liable to blame when they act partially would appeal, not to the moral purity of their intentions, but to the fact that they are engaged in what Derek Parfit calls "blameless wrongdoing" (*Reasons and Persons*, 32). The idea would be to claim that, in acting on loving dispositions toward their children, parents are acting on a set of dispositions that it was not wrong for them to acquire, even if on particular occasions (as when their bedtime-story reading leads them to make things

My first objection to the equality-weakly-first strategy raised doubts about the basis for the claim that well-off parents are excused, at the bar of justice, for bedtime-story reading. My second objection concerns the implications of claiming that they are so excused. Suppose, then, that they are. Parental bedtime-story reading, like parenting more generally, is an ongoing, repeated activity. Given this fact, it is questionable whether this view is consistent with parents maintaining, or reasonably being expected to maintain, the attitudes required to produce FRGs for and with their children. On Segall's view, such parents could only engage in bedtime-story reading knowing that what they are doing is an excusable violation of what justice requires of them. A well-off parent, sitting down to another night of bedtime-story reading, would apparently have to think: "I know that I am about to do something unjust. But never mind—I shall have an excuse."³⁵ Even if repeatedly engaging in preemptively excused activity of this kind is morally coherent, it is dubious whether it is psychologically coherent. And even if it is psychologically coherent, it is dubious whether it is consistent with maintaining—or reasonably being expected to maintain—the kinds of attitudes toward one's children and their interests that adequate parenting requires.³⁶ Whether it is in fact consistent with their maintaining such beliefs is an empirical question about which, I believe, we have little data, since I take it most well-off parents do not represent to themselves actions of the kind in question as unjust. It is not entirely an empirical question whether it would be reasonable to expect parents to maintain the requisite attitudes even if they knew that what they did was unjust, but it is enough for my purposes to note that the knowledge that what one is doing is unjust is no small thing. It might be that some parents could, in its face, maintain the attitudes

in a morally significant sense worse) their acting on these dispositions is impermissible. I cannot fully evaluate this intriguing possibility here. It will be enough for my purposes here to note that Parfit's arguments that these motivations are not wrong to acquire rest on a comparison between a world in which *no one* has partial motivations and a world in which most or all do; it is not clear that similar arguments would vindicate well-off people, in particular, acquiring such motivations.

- 35 One might reply that it would be better, then, that people did not generally *know* the basis on which they were not subject to coercive or other interference when they engage in bedtime-story reading. One might countenance, in other words, the relevant principle of equality being at least partly esoteric (cf. Glover and Scott-Taggart, "It Makes No Difference Whether or Not I Do It," 188–89). However, arguably it is a condition of adequacy on a principle of social justice that it not be in this way nonpublic (cf. Rawls, *A Theory of Justice*).
- 36 This second criticism arises with particular poignancy for Segall, who thinks it important that moral theory not issue guidance to agents that is morally and psychologically incoherent ("If You're a Luck Egalitarian, How Come You Read Bedtime Stories to Your Children?" 37).

necessary to produce FRGs; whether it would be reasonable to expect them to do so, in the face of the psychological cost and difficulty of so doing, however, is questionable. So I conclude that it is dubious whether, if parents have latitude to engage in bedtime-story reading only in the weak sense countenanced by this version of the equality-first strategy, we could recover our commonsense judgments about the kind of latitude that parents must have.

I have suggested that putting equality only weakly first does not successfully resolve liberal egalitarians' parental partiality puzzle. If we wish to reconcile equality and partiality, we so far seem to be left with only the option of putting partiality first. But, as noted at the end of the previous section, this strategy may seem unsatisfactory relative to the commitments of liberal egalitarianism. We can next ask, then, whether there is a well-motivated basis for putting equality first in a stronger sense.

3.2. *Putting Equality First in a Stronger Sense*

The trouble with trying to put equality first in a strong sense is that it is hard to see how this could be consistent with parents having latitude to engage in activities like bedtime-story reading. To illustrate the difficulty, consider the following case. Suppose you can give some significant gift to, or prevent some significant harm from befalling, your friend. And suppose that doing this will, as you know, unavoidably inflict significant suffering on some unrelated third party, although steps have been taken to minimize the degree of suffering that giving the benefit to your friend will impose on this third party. Translated into the terms of this simple case, the dispute between the partiality-first and equality-weakly-first strategies is about the correct weighing of these benefits and harms. The partiality-first strategy corresponds to the view that it matters more to give the benefit to one's friend, and hence that doing so is fully morally justified and so permitted *simpliciter*. The equality-weakly-first strategy corresponds to the view that the harm is significant enough that giving the gift to one's friend can be at most excused. And the simple case makes vivid how it may seem that there is no plausible way of reconciling equality and partiality by putting equality first in any stronger sense than this. That would be analogous to claiming that avoiding the harm to the third party matters much more than giving the gift to one's friend. It would seem to entail that, faced with the unavoidable conflict between benefiting one's friend or avoiding the harm to the third party, it is impermissible *simpliciter* to benefit one's friend. Translated back from the terms of the simple case, this would amount to a flat rejection of the latitude that commonsense morality affords to parental partiality. Putting equality first in any stronger sense, then, seems to amount to giving up the reconciliation project entirely.

The foregoing case against a strategy that puts equality first in a stronger sense, however, depended crucially on the assumption that well-off parents faced an *unavoidable* conflict—between, on the one hand, fulfilling their parental duties and, on the other, promoting equality. If this were so, then the partiality-first strategy would not, I believe, lack plausibility. Given only those two options, it would not be implausible to think it permissible (because morally justified) to fulfill one's parental duties even at the expense of not promoting equality. In the face of such an unavoidable conflict, it does not seem implausible to believe that equality ought to be subordinated to partiality. But the assumption that this conflict is unavoidable is questionable. It is an artifact of considering people's choices at a particular slice of time. When the assumption is dropped, so I shall next suggest, we arrive at a different way of reconciling equality and partiality that deserves to be taken seriously by liberal egalitarians.

Well-off people do not face an unavoidable choice between conflicting duties to respect partiality and promote equality. They have another option, one that would avoid their facing a choice in which they would inflict harms on some child or other—their own (in not producing FRGS) or someone else's (in undermining equal opportunity)—no matter what they do. They could simply refrain from having children. This opens the door to a different way of resolving the puzzle. The core idea behind this strategy is that well-off people are not permitted to parent. The partiality-first strategy reconciles equality and partiality in a way that grants permissions for *every person* to engage in a *more restricted* class of partial parental activities. The strategy now under consideration, by contrast, reconciles equality and partiality in a way that grants permissions for *only some persons* to engage in a *less restricted* class of partial parental activities. All and only those people of whom it is true that their acting partially toward their own children (consistent with background requirements of justice) would *not* disrupt equality would be licensed to act partially toward their children. In this sense, the strategy makes equality set the limits for when partiality is permissible. I shall therefore call this the strategy that puts equality first in a stronger sense.

The equality-strongly-first strategy solves the puzzle by rejecting P2. But it does not do so on a basis that flatly rejects the deliverances of commonsense morality; it is therefore still a reconciliation proposal. It retains a commitment to the idea that there must be permissions of partiality *for all those who can permissibly engage in* parenting. What it denies is that sufficiently well-off persons can permissibly do so. So it does not put anyone in the situation of having to choose between fulfilling their parental duties and promoting (or not disrupting) equal opportunity. Anyone can avoid putting themselves in such a position by ensuring they choose at most two from among the following three options:

(i) value equality, (ii) have children, or (iii) be comparatively well-off. You can either retain a commitment to equal opportunity and have children, so long as you forgo being comparatively well-off, or you can retain a commitment to equal opportunity and being well-off, so long as you forgo having children. Or you can give up a commitment to equal opportunity and indulge both a desire to be well-off and have children. But not all three. In short, if justice requires that people have a commitment to equal opportunity, then, other things fixed, the sufficiently well-off would not have children.³⁷

The equality-strongly-first strategy may initially seem incredible. It may seem to have the air of a (Jonathan) Swift-like modest proposal: liberal egalitarianism is to be made compatible with parental partiality by prohibiting the well-off from parenting—and so, other things fixed, prohibiting them from procreating at all. What could explain how this is a morally acceptable solution? I offer three reasons why the view deserves to be taken seriously. First, it is consistent with the motivations for liberal egalitarianism. Liberal egalitarianism's commitment to equal opportunity is motivated by the ideal of being taste-sensitive but circumstance-insensitive.³⁸ In other words, justice requires that how people fare be independent of their (unchosen) circumstances, but justice permits how they fare to be sensitive to tastes for which they are appropriately held responsible. In particular, how people fare can be sensitive to expensive tastes for which they can appropriately be held responsible. If someone forms or retains a preference for rare plover eggs in a way for which it is appropriate to hold them responsible, then justice does not require imposing costs on others, or redirecting resources from others, to subsidize the satisfaction of that taste.³⁹ But parenting by the well-to-do is a kind of expensive taste. Allowing a well-off person's taste for parenting to be satisfied, like subsidizing the satisfaction of someone's desire for rare plover eggs, imposes costs on others. Every person has an important interest in having access, on a basis they could reasonably accept, to the fruits of social cooperation; having this interest satisfied is a necessary condition of maintaining morally significant kinds of self-respect and a sense of one's self-worth.⁴⁰ Allowing a well-off person's taste for parenting to be satisfied therefore carries, for others, the significant cost of frustrating this interest. If it can be reasonable to hold well-off people responsible for forming or retaining a taste for parenting, and if justice is taste-sensitive, then there is a

37 See the next section for an explanation of this qualification.

38 For discussion and defense, see Cohen, "On the Currency of Egalitarian Justice," 916.

39 Cf. Cohen, "On the Currency of Egalitarian Justice," 913.

40 For discussion and defense of this claim, see, for example, Rawls, *Justice as Fairness*, 200.

principled basis within liberal egalitarianism for the view that justice disfavors subsidizing this taste and permitting well-off people to parent.⁴¹

Second, the degree of intuitive support that parenting by well-off people can claim from the precepts of commonsense morality is questionable. Suppose that some person had a share of the social product that was inconsistent with others having equal opportunity. Other things fixed, it would not be permissible for this person to keep this share; justice would require that it be taxed away so as not to upset the ideal of equal opportunity. But if this person is permitted to add a new person to the population and use his or her share of the social product to put that child in a position that is inconsistent with others having equal opportunity, then he or she could avoid what would otherwise have been required with respect to this share. Viewed in this light, parenting by the well-off seems to be an objectionable kind of self-dealing.⁴² And it is questionable how much intuitive support a social arrangement that permits such self-dealing can call on from the precepts of commonsense morality.⁴³

Third, although the strategy under consideration subordinates partiality to equality, it need not thereby fail to *respect* the demands of partiality. That is because partiality has the following peculiar feature: it is not obviously the kind of value that is to be *promoted*. What matters, one might claim, is to respect the ties of partiality that there are, not to create as many such ties as possible. (Compare: you have no partiality-given reason to make a booking now for a restaurant date in a year's time, even on the assumption you will meet a friend in the meantime whom you will have reason to take there. Doing so seems to betray a misunderstanding of the kind of value that partiality is.) But, as discussed above, the strategy presently under consideration subordinates partiality to equality, not by demanding that people fail to fulfill their partiality-given duties,

41 I return to this crucial question at the end of this section.

42 This is, in a sense, the inverse of a point made by Nozick; see *Anarchy, State, and Utopia*, 167–68. Nozick asks rhetorically how it could be that a patterned distributive principle could permit people to keep their shares for themselves but prohibit them from transferring their shares in certain ways to others. I am asking—also rhetorically but contrariwise—how it could be that a patterned distributive principle could be such that it does *not* permit someone to keep a certain share for herself but *does* permit her to transfer that share to someone else (of her choosing). But whereas Nozick's question is intended to elicit the judgment that a distributive principle must permit people to do as they please with the share the principle permits them to have, my question is intended to elicit the judgment that a distributive principle must not permit people to do as they please with a share that the principle would otherwise *not* permit them to have.

43 Of course, this person's child could, in turn, avoid the requirement to have their advantaged position taxed away only by having children. So perhaps the more apt comparison is with a Ponzi scheme, rather than a self-dealing scheme.

but rather by demanding only that certain people not form certain relationships with persons with whom they would then *have* certain partiality-given duties. So the strategy in question does not obviously result in any morally significant compromise of what partiality requires.

That concludes my *prima facie* case that the equality-strongly-first strategy deserves to be taken seriously by liberal egalitarians. Could this strategy succeed in solving the puzzle? In answering this question, the crucial question for the liberal egalitarian is whether or not people's interest in parenting is so weighty as to justify subsidizing this taste when it is—as it is in the case of wealthy people—an expensive taste, one that imposes serious costs on others. That is the crux of the dispute between the partiality-first and equality-strongly-first strategies. Against the equality-strongly-first strategy, it might be claimed that the interest that any person (including the wealthy) has in parenting is weighty enough to make it *not* reasonable to hold people responsible for forming or retaining this taste. And this claim would not be obviously mistaken. For some people, playing the fiduciary role of parent is a central, ineliminable element of their living a life in which they flourish.⁴⁴ The difficult question, then, is how to weigh this interest against the interest—grounded in self-respect—in living in a society in which one's access to positions of advantage is independent of the unchosen circumstances of one's birth.⁴⁵ It is not, I believe, at all clear how to weigh these quite different interests. I do not attempt to do so here. We thus, I conclude, have on the table two views that are well-positioned to resolve the parental partiality puzzle for liberal egalitarianism: the partiality-first strategy and the equality-strongly-first strategy. Which of these two views liberal egalitarians have most reason to adopt, I have suggested, depends on the difficult question just mentioned about weighing the relative moral importance of two quite different interests that people have.

4. OBJECTIONS

I have given a *prima facie* case that liberal egalitarians should take seriously the strategy of putting equality first in a stronger sense. But it might be thought that this *prima facie* case does not stack up on reflection. In this section, I consider two objections that might be thought to defeat the *prima facie* case that I have offered for putting equality strongly first. Each objection suggests that the strategy has repugnant or counterintuitive implications in practice.

44 See, for discussion and defense, Brighouse and Swift, "Legitimate Parental Partiality," 54.

45 For arguments that suggest that this interest is grounded in self-respect, see, for example, Rawls, *Justice as Fairness*, 200.

First, one might think that the equality-strongly-first strategy implies that those who are now rich and now have children should have them forcibly taken away. But, in reply, it does not have this or any other implausibly draconian implication for public policy. The question on the table is about the basic attitude that it is appropriate to take toward the prospect of our social arrangements permitting well-off people to engage in a certain class of partial activities. The answer to this question about a fundamental principle of justice is distantly removed from the question of what social rules of regulation we are to have when we take into account not only considerations of justice but a host of other noninstrumental and instrumental reasons that bear on this issue. There might, for example, be noninstrumental reasons of solidarity to permit everyone to engage in every activity that is legally permitted. There might, for example, be instrumental reasons of stability to permit everyone to engage in an activity in which they have a very strong, identity-involving desire to engage. All this is downstream of the question presently on the table: Supposing that well-off people engage in parenting, is there an injustice? It is only that question to which the equality-strongly-first strategy gives an (affirmative) answer.⁴⁶ For similar reasons, the strategy does not entail that well-off persons as a group are blameworthy for engaging in parenting, nor that some individual well-off person is blameworthy for doing so given that many other well-off persons are already doing so in our present social circumstances.

Nor would the equality-strongly-first strategy *entirely* foreclose on well-off persons parenting without injustice. Recall that its claim is only that, other things fixed, in a just society the well-off would not parent. But other things might not be fixed. To see how this is possible, consider the following comparison. Effective altruists need not spend their whole lives working directly for the badly off. They can instead earn a lot of money, for example, on Wall Street, and “export” their duties to benefit the badly off by using their earnings to subsidize others who can do more, more effectively, than they can for worse-off people. Likewise, prospective parents could indulge their desires to be comparatively well-off and raise children, consistent with a commitment to equality, by “exporting” to others their duties to promote equal opportunity. They might, for example, raise children who are themselves strict and effective equality promoters. If those children could do more to promote equality than those parents could have done by refraining from having children (or if parents

46 Indeed, the equality-strongly-first strategy is consistent with accepting the *content* of Brighthouse and Swift’s partiality-first strategy or Segall’s equality-weakly-first strategy, reinterpreted as views about the “contrary-to-duty” or “nonideal” case in which the well-off choose to parent, ignoring what (by the lights of the equality-strongly-first strategy) they have most reason of justice to do.

could at least reasonably believe this to be so), then they could, consistent with the equality-strongly-first strategy, permissibly engage in parenting. In so doing they would not, on balance, be making the idea of equal opportunity less well-realized.⁴⁷

Consider next a different objection. I have characterized the equality-strongly-first strategy as (other things fixed) prohibiting the *well-off* from parenting. But there is significant empirical evidence that, in our social circumstances, (a) parenting styles are a strong predictor of children's socioeconomic status and (b) a "concerted cultivation" parenting style (including, e.g., the reading of bedtime stories) is common not only to the *well-off* but to every social class *except* for those who are quite badly off.⁴⁸ But if parenting style is the mechanism by which parental advantage is reproduced in children, and if the relevant advantage-conferring parenting style is common to all but those who are badly off, then the equality-strongly-first strategy would seem to entail that it is unjust, not only for the *well-off* to parent, but for *anyone* above the average level of opportunity to parent. Such a radical revision of commonsense beliefs threatens the plausibility of solving the parental partiality puzzle by putting equality strongly first.⁴⁹

This objection raises a serious challenge to the equality-strongly-first strategy. The underlying fact to which it draws attention is this: how far-reaching the practical upshot of the strategy will be depends in part on the relative

47 It might be, however, that such parenting, which would "enroll" children in a view about what matters in life, runs afoul of independent constraints on raising children in ways to which they could reasonably consent (cf. Clayton, "Debate"). Arguably, however, "enrolling" children in a commitment to *reasonable* values does not run afoul of this constraint (cf. Cormier, "On the Permissibility of Shaping Children's Values"). I leave aside here the interesting question of whether or not raising children to be equal-opportunity promoters is an objectionable kind of "enrollment" in this sense.

48 For discussion and defense, see further Lareau, *Unequal Childhoods*, 1–33, 233–39. More particularly, Lareau characterizes a "concerted cultivation" parenting style as follows:

Parents actively fostered and assessed their children's talents, opinions, and skills. They scheduled their children for activities. They reasoned with them. They hovered over them and outside the home. They did not hesitate to intervene on the children's behalf. They made a deliberate and sustained effort to stimulate children's development and to cultivate their cognitive and social skills. (*Unequal Childhoods*, 238)

The contrast is with a parenting style that emphasizes spontaneity and "natural growth" (Lareau, *Unequal Childhoods*, 238) that is associated with working-class families who lack the time and/or monetary resources for concerted cultivation. For discussion of how Lareau's work bears on the tension between equality and partiality, see Brighouse and Swift, *Family Values*, 127.

49 I thank an anonymous reviewer for the journal for suggesting this objection.

proportions of people in different social classes. In highly stratified societies (i.e., a society in which a small number are far above the average level of opportunity and most are significantly below it), putting equality strongly first will imply that only a comparatively small number of people cannot permissibly parent. In societies in which a larger number of people are only moderately above the average level of opportunity (and in which conditions a and b above hold), putting equality strongly first will cast a wider net. To the extent that it is a moral datum that, in *any* society (no matter the relative numbers in different social classes), considerations of justice could not oppose *most* people's parenting, this is an undeniable cost of putting equality strongly first.

Even if the above is a moral datum, however, the objection may not be a decisive strike against putting equality strongly first. The force of the objections depends, at least in part, on how the badness of inequality of opportunity is determined. To see this, consider the following illustrative case. Suppose there are three groups—*R*, *M*, and *P*—such that *R* and *P* are respectively far above and far below the average level of opportunity and *M* is slightly above it. Now suppose that many people in *M* choose to parent and that they reliably transmit their level of advantage to their children via a concerted-cultivation-like style of parenting. Grant that, on the equality-strongly-first strategy, there would be a reason of justice disfavoring these people's choice. We can then ask: How seriously would these people thereby have disrupted equal opportunity? The answer depends on how to measure the badness of unequal opportunity. Using the useful framework developed by Larry Temkin, we can consider how bad a situation is with respect to unequal opportunity as a function of the *complaints* held by each person with respect to unequal opportunity.⁵⁰ The overall badness of the situation then depends on who has a complaint, how serious is the complaint, and how are complaints aggregated. Consider, for example, the following pair of answers:

1. (a) Everyone worse off than the highest level of opportunity has a complaint, (b) each complainant has a complaint relative to all those better off in opportunities than her, and (c) complaints are aggregated by weighting more heavily those with lower levels of opportunity.
2. (a) Everyone with a level of opportunity below the average level has a complaint, (b) each complainant has a complaint relative to all those at the highest levels of opportunity, and (c) complaints are aggregated by simple addition.

50 Temkin, *Inequality*, ch. 2.

These yield very different answers to the question about our simple case. Suppose a large number of people in *M* choose to parent. If 1 is correct, then (because of 1a and 1b) each child in *P* would have a large number of new complaints, each one of significant size; and (because of 1c) each new complaint would contribute significantly to the overall outcome's overall unequal-opportunity "score." On this view, parenting by those in *M* would significantly worsen unequal opportunity. But if instead 2 is correct, then assuming that more *M*-children does not substantially affect the average level of opportunity enjoyed, then (because of 2b) children in *P* would not have significant new complaints, nor (because of 2a) would the extra children in *M*, and (because of 2c) even if complaints among *P*-children have increased in size to some degree, they would not disproportionately affect the outcome's overall unequal-opportunity "score." So, on this view, parenting by those in *M* need not significantly worsen unequal opportunity. Thus even if there is a reason of justice that disfavors people in *M* choosing to parent, on this view it would be significantly less weighty than the reason of justice that disfavors people in *R* parenting. And this could be true even if there are many more people in *M* than in *R*. The upshot is that the equality-strongly-first strategy need not entail that, whenever some people's equal-opportunity interests come into unavoidable conflict with other people's interest in parenting, the former must be weightier. Given an appropriate measure of the badness of unequal opportunity, the strategy is consistent with only parenting by the quite well-off being disfavored by a *weighty* reason of justice. Since it is unsettled how to measure the badness of unequal opportunity, I therefore conclude that the objection may not be a decisive strike against the equality-strongly-first strategy.⁵¹

5. CONCLUSION

Let me briefly summarize the discussion. I have been considering a puzzle for liberal egalitarians that arises from the fact there appears to be a conflict between their commitment to a robust ideal of equality and the latitude that commonsense morality gives to parental partiality (section 1). In resolving the puzzle, it is natural to begin from the attractive idea of trying to find a basis for reducing the apparent conflict by licensing only parental partial activities (like bedtime-story reading) that are necessary to produce a morally significant kind of good and then putting partiality first in these cases of putatively

51 In responding to this objection, the proponent of putting equality strongly first could also appeal to the point made in response to the earlier objection. It may be easier for those whose parenting would not result in further significant dis-equalization of opportunity to "offset" the effect of indulging in their (only somewhat) expensive taste for parenting.

unavoidable conflict. But, so I suggested, it is questionable whether this strategy accords well with the importance liberal egalitarians wish to give to equality (section 2). And we could not, I then suggested, plausibly resolve the tension between equality and partiality in cases of putatively unavoidable conflict by putting equality only weakly first, claiming that parents are merely excused when they engage in partial activities (section 3.1). We could, however, plausibly put equality first in a stronger sense, by denying that the well-off really face an unavoidable conflict between discharging parental duties and promoting equality (section 3.2).

We are thus left, I concluded, with two ways of resolving the parental partiality puzzle for liberal egalitarians: the partiality-first strategy, corresponding to the claim that equality matters less than partiality, or an equality-first strategy, corresponding to the claim that partiality matters less than equality. The former strategy recovers a narrower range of permissions for a wider class of persons to engage in parenting; the latter recovers a wider range of permissions for a narrower class of persons to engage in parenting. Both strategies, I have suggested, are consistent with the motivations for liberal egalitarianism. Which one is more plausible depends on settling the question of whether people's interest in parenting is more or less weighty than people's interest in there being equality of opportunity—and may depend on how to measure the badness of unequal opportunity (section 4).⁵²

Tulane University
dobriemo@tulane.edu

REFERENCES

- Arneson, Richard. "Justice Is Not Equality." *Ratio* 21, no. 4 (December 2008): 371–91.
- Bou-Habib, Paul. "The Moralized View of Parental Partiality." *Journal of Political Philosophy* 22, no. 1 (March 2014): 66–83.
- Brake, Elizabeth. "Fair Care: Elder Care and Distributive Justice." *Politics, Philosophy, and Economics* 16, no. 2 (May 2017): 132–51.

52 I am deeply grateful to Anca Gheaus, Gina Schouten, and Harry Brighouse for written comments on and discussion of earlier drafts of this paper. Special thanks are also due to the remarkable students in my fall 2019 Social and Political Ethics class at Tulane University—especially Talia Boninger, Alex Calloway, Madeline Cargill, Fiona Hellerman, Max Jaffe, Ann Kapustiak, and Alec Wild—for the lively discussion that originally prompted the paper.

- Brighouse, Harry, and Adam Swift. *Family Values: The Ethics of Parent-Child Relationships*. Princeton, NJ: Princeton University Press, 2014.
- . “Legitimate Parental Partiality.” *Philosophy and Public Affairs* 37, no. 1 (Winter 2009): 43–80.
- Brownlee, Kimberley. “A Human Right against Social Deprivation.” *Philosophical Quarterly* 63, no. 251 (April 2013): 199–222.
- Clayton, Matthew. “Debate: The Case against the Comprehensive Enrolment of Children.” *Journal of Political Philosophy* 20, no. 3 (September 2012): 353–64.
- Cohen, Gerald Allan. “On the Currency of Egalitarian Justice.” *Ethics* 99, no. 4 (July 1989): 906–44.
- . *Rescuing Justice and Equality*. Cambridge, MA: Harvard University Press, 2008.
- . *Why Not Socialism?* Princeton, NJ: Princeton University Press, 2009.
- Cordelli, Chiara. “Distributive Justice and the Problem of Friendship.” *Political Studies* 63, no. 3 (August 2015): 679–95.
- Cormier, Andree-Anne. “On the Permissibility of Shaping Children’s Values.” *Critical Review of International Social and Political Philosophy* 21, no. 3 (2018): 333–50.
- Enoch, David. “The Masses and the Elites: Political Philosophy for the Age of Brexit, Trump, and Netanyahu.” *Jurisprudence* 8, no. 1 (March 2017): 1–22.
- Foot, Philippa. “The Problem of Abortion and the Doctrine of Double Effect.” *Oxford Review* 5 (1967): 1–7.
- Freeman, Samuel. “The Basic Structure of Society as the Primary Subject of Justice.” In *A Companion to Rawls*, edited by Jon Mandle and David A. Reidy, 88–111. Oxford: Wiley-Blackwell, 2013.
- Gheaus, Anca. “How Much of What Matters Can We Redistribute? Love, Justice, and Luck.” *Hypatia* 24, no. 4 (Fall 2009): 68–90.
- . “Love and Justice: A Paradox?” *Canadian Journal of Philosophy* 47, no. 6 (April 2017): 739–59.
- . “Personal Relationship Goods.” *Stanford Encyclopedia of Philosophy* (Fall 2018). <https://plato.stanford.edu/archives/fall2018/entries/personal-relationship-goods>.
- Glover, Jonathan, and M. J. Scott-Taggart. “It Makes No Difference Whether or Not I Do It.” *Proceedings of the Aristotelian Society* 49 (1975): 171–209.
- Lareau, Annette. *Unequal Childhoods: Class, Race, and Family Life*. Berkeley: University of California Press, 2003.
- Macleod, Colin. “Liberal Equality and the Affective Family.” In *The Moral and Political Status of Children*, edited by David Archard and Colin Macleod, 212–30. Oxford: Oxford University Press, 2002.

- . “The Puzzle of Parental Partiality: Reflections on *How Not to Be a Hypocrite: School Choice for the Morally Perplexed Parent*.” *Theory and Research in Education* 2, no. 3 (November 2004): 309–21.
- Munoz-Dardé, Veronique. “Is the Family to Be Abolished Then?” *Proceedings of the Aristotelian Society* 99 (1999): 37–56.
- Neufeld, Blain. “Coercion, the Basic Structure, and the Family.” *Journal of Social Philosophy* 40, no. 1 (Spring 2009): 37–54.
- Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.
- Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Pogge, Thomas. “On the Site of Distributive Justice: Reflections on Cohen and Murphy.” *Philosophy and Public Affairs* 29, no. 2 (April 2000): 137–69.
- Rawls, John. *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press, 2001.
- . *A Theory of Justice*. Rev. ed. Cambridge, MA: Harvard University Press, 1999.
- Scheffler, Samuel. *Equality and Tradition*. New York: Oxford University Press, 2012.
- Schouten, Gina. “Restricting Justice: Political Interventions in the Home and the Market.” *Philosophy and Public Affairs* 41, no. 4 (Fall 2013): 357–88.
- Segall, Shlomi. “If You’re a Luck Egalitarian, How Come You Read Bedtime Stories to Your Children?” *Critical Review of International Social and Political Philosophy* 14, no. 1 (2011): 23–40.
- Swift, Adam. *How Not to Be a Hypocrite: School Choice for the Morally Perplexed Parent*. London: Routledge, 2003.
- Temkin, Larry. *Inequality*. Oxford: Oxford University Press, 1993.
- . “Justice, Equality, Fairness, Desert, Rights, Free Will, Responsibility, and Luck.” In *Responsibility and Distributive Justice*, edited by Matthew Clayton and Zofia Stemplowska, 51–76. Oxford: Oxford University Press, 2011.

DEMOCRACY AND SOCIAL EQUALITY

Ryan Cox

RECENT egalitarian theorizing has drawn attention to the importance of a conception of social equality understood in terms of all members of a society relating to each other as equals.¹ On such a conception of social equality, there appears to be a particularly intimate connection between democracy and social equality. Democracy, understood in terms of all members of a society having an equal opportunity to influence political decision-making, appears to be *necessary* for social equality, understood in terms of all members of a society relating to each other as equals. Democracy appears to be an important *constituent* of social equality.

While egalitarian theorists have often assumed that there is an intimate relation between democracy and social equality, only recently have they attempted to clarify the exact nature of the relation between social equality and democracy and to argue that such a relation obtains. Earlier egalitarian theorists seem to have held that while social equality was intimately related to democracy, democracy only bore a *contingent* and *causal* relation to social equality—democracy, perhaps, was seen to be more likely than other forms of government to lead to social equality. Many recent egalitarian theorists, however, have wanted to argue that the relation between democracy and social equality is more intimate than this. They have wanted to argue that democracy is *necessary* for social equality and that it is an important *constituent* of social equality. For instance, Niko Kolodny has argued that democracy is “necessary for full or ideal social equality” and that it “is a particularly important constituent of a society in which people are related to one another as social equals.”² Elizabeth Anderson has argued that hierarchies of command are constitutive

1 See, for instance, Wolff, “Fairness, Respect, and the Egalitarian Ethos” and “Equality and Hierarchy”; Anderson, “What Is the Point of Equality?” and “Equality”; Scheffler, “What Is Egalitarianism?” and “The Practice of Equality”; Kolodny, “Rule over None I” and “Rule over None II.” For work on social equality prior to this cluster, see note 12. See Fourie, Schuppert, and Wallimann-Helmer, *Social Equality*, for a collection of recent work on social equality.

2 Kolodny, “Rule over None II,” 308, and “Rule over None I,” 196.

of social inequality and has investigated the ways in which democratic forms of government turn relations of rulers and subjects into relations of agents to principals, where “democratically accountable officeholders” do not “constitute a social hierarchy in the sense of distinct classes of rulers and ruled.”³ Samuel Scheffler has argued that “the ideal of a society of equals” is “subject to the presumption that each participant in an egalitarian relationship is equally entitled to participate in decisions made within the context of the relationship.”⁴ And, more recently, Daniel Viehoff has written that he is “sympathetic to the theory that equality is a constitutive component of certain non-derivatively valuable relationships.”⁵

The question of the nature of the relation between democracy and social equality bears directly on the issue of the nature of justifications of democracy in terms of social equality. Many democratic theorists seek noninstrumental justifications of democracy *over* alternative forms of government.⁶ If democracy—understood as requiring political equality—is necessary for full social equality, then insofar as we have an interest in relating as equals to those with whom we stand in ongoing social relations, we will have a noninstrumental *pro tanto* reason to support democracy over alternative forms of government. Thus, we have the beginnings of the kind of justification of democracy that goes beyond the standard instrumental justifications of democracy and may favor democracy over the alternatives, even if democracy does not turn out to be the instrumentally best form of government.⁷

My aim in this essay will be to argue that the relation between democracy and social equality is, at best, a *contingent* and *causal* relation. It follows that there can be no noninstrumental justification of democracy in terms of social inequality, so understood.⁸ While I agree with egalitarian theorists that we

3 Anderson, “Equality,” 43–44, 48.

4 Scheffler, “The Practice of Equality,” 38.

5 Viehoff, “Power and Equality,” 8.

6 For discussions of the structure of such justifications, see Wall, “Democracy and Equality”; and Viehoff, “Power and Equality,” 5.

7 According to Thomas Christiano and John Christman, the debate over “whether democracy is merely instrumentally justified or whether there is some intrinsic merit to democratic ways of making decisions” is one of the “two main sources of debate concerning the normative underpinnings of democracy” (*Contemporary Debates in Political Philosophy*, 10).

8 I am not claiming that there can be no noninstrumental justification of democracy at all. Thomas Christiano’s account of the relation between democracy and equality may avoid my criticisms since Christiano is not operating with an understanding of social equality cast in terms of the notion of relating as equals. Christiano derives his understanding of social equality directly from a deeper understanding of moral equality understood in terms of equal consideration of interests. See Christiano, *The Constitution of Equality*.

have an interest in relating as equals to those with whom we stand in ongoing social relations—that is, I agree with the egalitarian assumption behind this kind of argument for democracy—and while I agree that there is an intimate relation between democracy and social equality, I believe that it is a mistake to think, as recent egalitarian theorists do, that democracy is *necessary* for social equality and that democracy is an important *constituent* of social equality. My case against the constitution thesis—the thesis that democracy is necessary for social equality and is an important constituent of social equality—will involve presenting a series of cases that show how democracy and social equality can come apart. This, of course, is an appropriate methodology for assessing claims of necessity and constitution. However, I will not rest my case entirely on such judgments about these cases. By presenting an alternative account of the relation between democracy and social equality, one on which democracy is merely contingently and causally related to social equality, I will attempt to explain away the initial appearance of a necessary and constitutive connection between democracy and social equality. I will argue that a better explanation of what is problematic about inequalities in power and *de facto* authority in the political context is that they either can be *used* to establish or reinforce social inequalities through the direct exercise of that power or can be used by those with perceived greater power and *de facto* authority to extract greater consideration from those with less power and authority.⁹

The discussion is structured as follows. I begin, in section 1, by locating the topic of social equality and examining the argument for thinking that democracy is an important constituent of social equality and is necessary for full social equality. In section 2, I argue on the basis of a series of examples against the claim that democracy is necessary for full social equality. I argue that these examples show that the relation between democracy and social equality is, at best, causal and contingent, and I use the examples to motivate the alternative extraction theory. In section 3, I compare my argument for this conclusion with a related argument by Richard Arneson. In section 4, I briefly discuss the implications of my conclusion for noninstrumental justifications of democracy.

9 This view thus has a similar structure to the theory of justice and complex equality defended by Michael Walzer. But whereas Walzer thinks that political power can easily be used to establish and maintain inequalities in particular distributive spheres, and so might upset complex equality, I am claiming that it can easily be used to establish and maintain social inequalities (Walzer, *Spheres of Justice*, 15). Such a view could be thought of as the natural result of combining Walzer's claims in this context with David Miller's elaboration of Walzer's position. See Miller, "Complex Equality." The view is also similar to Philip Pettit's view of the relation between domination and social inequality. I discuss Pettit's position in section 3.

1. DEMOCRACY AND SOCIAL EQUALITY

I will begin, in this section, by locating the topic of social equality. It is important for my purposes that we have an independent, pretheoretical grasp of the notion of social equality and of its normative significance. I will then make some general remarks about the structure of claims about social equality and the normative foundations of social equality before turning to recent analyses of the notion of social equality. With these theoretical and pretheoretical understandings in hand, I will then examine the case for thinking that democracy is necessary for social equality and an important constituent of social equality.

1.1. *Social Equality: Locating the Topic*

The notion of social equality of interest to us is best introduced in terms of the notion of members of a society relating to each other as equals. Egalitarian theorists often speak of social equality in such terms. For instance, Elizabeth Anderson writes that egalitarians are committed “to creat[ing] a community in which people stand in relations of equality to others.”¹⁰ One way of understanding talk of “relations of equality” is in terms of *how members of a society relate to each other*. Do the members of a society relate to each other as equals or do they relate to each other as inferiors to superiors or superiors to inferiors? Standing in relations of equality to each other, on this understanding, is a matter of relating to each other as equals as opposed to superiors or inferiors. This understanding is in play when Samuel Scheffler writes that egalitarians should care about “the establishment of a society of equals, whose members relate to one another on a footing of equality.”¹¹ This understanding is in play when David Miller writes:

[Equality] identifies a social ideal, the ideal of a society in which people regard and treat one another as equals—in other words, a society that does not place people in hierarchically ranked categories such as classes. We can call this . . . kind of equality *equality of status* or simply *social equality*.¹²

10 Anderson, “What Is the Point of Equality?” 289.

11 Scheffler, “The Practice of Equality,” 21.

12 Miller, *Principles of Social Justice*, 232. The understanding of social equality in terms of “relating as equals” is prevalent in the egalitarian literature. Miller writes that equality “identifies a form of life in which people in a very important sense treat one another as equals. In their social intercourse, they act on the assumption that each person has an equal standing that transcends particular inequalities (of achievement, for instance)” (*Principles of Social Justice*, 240). See also Miller, “Equality and Market Socialism,” 301–2. The preface of Walzer’s *Spheres of Justice* can be read as a tour de force statement of this ideal of social equality.

I assume that we each have a good pretheoretical understanding of the notion of relating to each other as equals involved in social equality. We are able to identify paradigmatic cases where some members of a society relate to each other as superiors to inferiors. We immediately recognize the relations of lord to peasant, of patricians to plebeians, of Brahmins to untouchables, and of Black people to White people (in the Reconstruction-era South, at least) as relations of superior to inferior. Not only do we immediately recognize these relations as relations of superior to inferior, but we also immediately recognize the normative significance of these relations: they are relations we have good reason to avoid. It is by reflecting on such paradigmatic examples, and variations on such examples, and using our pretheoretical understanding of the notion of relating to each other as equals that we can start to build a theory of social equality. What is it about these paradigmatic examples of social inequality that make them examples of social inequality?

We will examine an attempt to answer this question in a moment. But before doing so, I want to make a few remarks about the structure of claims about social equality that should be common ground among theories of social equality. At the heart of the notion of relating as equals is the notion of relating to each other in a society. The notion of relating to each other is a capacious one, and it includes all the ways that members of a society may relate to each other in a society, from how they interact personally to how they structure their physical environment. I take it to be a task of a theory of social equality to say which of these ways of relating are ways of relating as equals and which are ways of relating as superior to inferior. It is plausible that some ways of relating to each other will only count as ways of relating as superior to inferior in certain contexts or in certain circumstances. So, for instance, relating to another as one with greater power or authority may

Walzer writes: “The experience of subordination—of personal subordination, above all—lies behind the vision of equality” (xiii); “The aim of political egalitarianism is a society free from domination” (xiii); “No more bowing and scraping, fawning and toadying; no more fearful trembling; no more high-and-mightiness; no more masters, no more slaves” (xiii). Walzer speaks of “a certain conception of how human beings relate to one another and how they use the things they make to shape their relations” (xiv). Anderson writes, “To be an egalitarian is to commend and promote a society in which members interact as equals” (*Private Government*, 3). Richard Arneson, a particularly perceptive critic of relational egalitarianism rightly takes the idea that members of a society must “relate to each other as equals” to be central to the approach (“Democratic Equality and Relating as Equals”). Two classic works that gesture toward this conception of social equality as relating as equals are Crosland, *The Future of Socialism*; and Tawney, *Equality*. See Miller, “Equality and Market Socialism,” 300–3, for a brief discussion of these works. See also Benn and Peters, *Social Principles and the Democratic State*, 120–22. Anderson traces the ideal of social equal in this sense back to the Levellers in seventeenth-century England. See Anderson, *Private Government*, 7–17.

not be a way of relating as superior to inferior in certain contexts, but it may be in another context. As we are about to see, recent egalitarian theorists hold that relating to another as one with greater power or authority *does* amount to relating as superior to inferior, at least in the political context.

As I just said, we immediately recognize the normative significance of relations of social inequality. We recognize that they are relations we have good reason to avoid. I agree with egalitarian theorists that the reasons we have to avoid such relations are noninstrumental. We have reasons to avoid such relations for their own sake. It may well be that the claim that we have good reason to avoid social inequality is moral bedrock. It may be the kind of claim that does not admit of further defense. I suspect, however, that the claim that we *ought* to *relate* to each other as equals is somehow grounded in the idea that we *are*, in some important sense, equals. We ought to relate to each other as the equals we in fact are.¹³ In the absence of now discredited views that some human beings *are* naturally superior or naturally inferior, most people now believe that all human beings *are* morally equal and, as such, ought to relate to each other as the equals we are. As a consequence, our social relations—how we relate to each other in society—ought to reflect this basic fact about us, and we ought to relate to each other as the equals we are. It is a familiar point, however, that relating to one another “as equals” may involve a great deal of unequal treatment. Our question is whether political equalities are the kind of equalities that are required for the kind of treatment as equals, relating to each other as equals, required for social equality. I mention this foundation of the normative significance of social equality only to point out that it is plausible, whether we take it to be moral bedrock or not, that we have good reason to avoid social inequality for its own sake.¹⁴

Having located the topic of social equality, let us now turn to recent attempts to answer the question of which ways of relating to each other in particular contexts or particular circumstances are ways of relating as equals and which are ways of relating as superior to inferior.

1.2. Analyses of Social Equality

Recent egalitarian theorists have attempted to analyze social equality and inequality in order to isolate those features of the paradigmatic cases that make them cases of social equality or social inequality. Perhaps the most sustained

13 See Williams, “The Idea of Equality,” for discussion of how descriptive and prescriptive claims of equality relate to each other. Dworkin, “What is Equality?” is the classic source of the distinction between treatment *as* equals and equal treatment.

14 Cf. “Why do socialists find social equality attractive? In the end, perhaps, this will have to be regarded simply as an ethical commitment, an ultimate goal not capable of justification to someone who does not sympathize with it” (Miller, “Equality and Market Socialism,” 302).

attempt to offer such an analysis in the recent literature comes from Niko Kolodny. Whereas theorists like Anderson have drawn attention to important differences between different kinds of social hierarchies and their characteristic varieties—hierarchies of domination and command, hierarchies of esteem, and hierarchies of standing—Kolodny has been the most explicit about attempting to isolate the features of these different kinds of hierarchies that *make them* problematic cases of social inequality.¹⁵ It is partly for this reason that I will focus on Kolodny’s analysis here. Another reason for focusing on Kolodny’s account is that, as we will see below, Kolodny foresees a particular kind of objection to analyses of social equality in terms of equalities of power and authority and modifies his account in order to avoid this kind of objection. Although I will mainly focus on Kolodny’s analysis, I take the arguments of the next section to tell against any attempt to analyze social equality in terms of equality in power and *de facto* authority.

As I mentioned before, we can begin to develop an account or analysis of social equality by reflecting on paradigmatic cases of social inequality, drawing on our pretheoretical understanding of relating as equals, and attempting to isolate the features that make them cases of social inequality. This is the approach taken by Kolodny. Since it is difficult to distinguish judgments about social equality, and relating as equals, from judgments about distributive equality or distributive justice, Kolodny considers an example of a “society administered by a class of ascetic warriors” in which distributive justice is perfectly administered and the warriors themselves in no way benefit from the distribution they administer: each is given their due, including the warriors themselves, and any benefits that accrue to them in virtue of having the authority they do is offset by their depriving themselves of other advantages. Kolodny writes that even in such a society, where everyone has been given their due, “there is an obvious sense in which [the ascetic warriors] constitute a superior social stratum [and] occupy a higher position in the hierarchy.”¹⁶ Reflecting on this example, Kolodny asks, “What is present in the societies that we have described . . . that might account for the intuitive presence of social inequality?”¹⁷ According to Kolodny, “it seems to have to do with the following”:

- (i) Some having greater relative *power* (whether formal or legal or otherwise) over others, while not being resolutely disposed to refrain

15 See Anderson, “What Is the Point of Equality?” and “Equality.”

16 Kolodny, “Rule over None II,” 295. I will question whether there is an “obvious sense” in which the ascetic warriors constitute a superior social stratum at the end of section 2. For the purpose of explication, I grant Kolodny the appearance of social inequality here.

17 Kolodny, “Rule over None II,” 295.

from exercising that greater power as something to which others are entitled.

- (ii) Some having greater relative *de facto* authority (whether formal or legal or otherwise) over others, in the sense that their commands or requests are generally, if not exceptionlessly, complied with (although not necessarily for any moral reasons), while not being resolutely disposed to refrain from exercising that greater authority as something to which those others are entitled.
- (iii) Some having attributes (e.g., race, lineage, wealth, perceived divine favor) that generally attract greater *consideration* than the corresponding attributes of others.¹⁸

These are plausible candidates for being the features of the example in virtue of which it is an example of social inequality.¹⁹ Since it will be important for the discussion to follow, I want to draw attention to the “while” clauses in (i) and (ii): “while not being resolutely disposed to refrain from exercising that greater power/authority as something to which others are entitled.” According to Kolodny, “What social equality requires is that ‘natural’ power be regulated by the *right* dispositions.”²⁰ The thought here is that some having greater relative power over others may not be seen as problematic if those who have the greater power are resolutely disposed to refrain from exercising that power as something to which others are entitled, for in such a case the inequality in power will be regulated by the right dispositions. It is only when those with greater power are not so disposed that greater power becomes problematic. As we will see, this qualification is essential to Kolodny’s view, and we will be able to distinguish a *simple* view of the relation between inequalities of power and *de facto* authority and social inequality (a view that takes such inequalities themselves to suffice for social inequality) and Kolodny’s more complex view (a view that holds that such inequalities in the absence of the relevant dispositions and in certain circumstances suffice for social inequality).

These are the three characteristic features of social inequality then: asymmetries in power, asymmetries in *de facto* authority, and asymmetries in

18 Kolodny, “Rule over None II,” 295–96.

19 Kolodny says that these three conditions select what he thinks are the essential features of the analysis of social inequality given by Anderson and Scheffler (Kolodny, “Rule over None II,” 296). I agree that these three conditions get to the heart of egalitarian concerns about these kinds of examples.

20 Kolodny, “Rule over None II,” 296n8.

consideration.²¹ These asymmetries are characteristic of cases where members of a society fail to relate to each other as equals. Now, we might ask whether each asymmetry is necessary for social inequality or whether any asymmetry is independently sufficient for it. Kolodny suggests that the consideration component of social equality “need not be essential to the argument that a concern for social equality implies a concern for democracy.”²² He thinks that the power and *de facto* authority components “may suffice for the argument.”²³ In other words, according to Kolodny, asymmetries in power and *de facto* authority—in certain contexts—might suffice for social inequality; that is, they might suffice for a failure to relate as equals. Of course, asymmetries of consideration are nonetheless central to our understanding of social inequality, and it seems that asymmetries here, while perhaps not necessary, may independently *suffice* for social inequality. This is a point to which we will return later.

1.3. *The Relation between Democracy and Social Equality*

Now that we have located the topic of social equality and have a candidate analysis of social inequality in hand, we can ask how democracy is supposed to relate to social equality so understood. As I said earlier, whether a way of relating to each other is a way of relating as equals or as superior to inferior is often a matter of context. Relating to another as one with greater power need not always amount to relating as superior to inferior. After all, there are inequalities in power and *de facto* authority across many contexts of our social lives, from churches to universities and workplaces, and we do not find these inequalities in power and authority problematic, and we do not think they suffice for a kind of social inequality.²⁴ In order to argue that democracy is *necessary* for social equality and an important *constituent* of it, we need to know what it is about the political context that explains why inequalities in power and *de facto* authority *suffice*, in this context, for relations of social inequality. According to Kolodny, the “three features of political decisions” relevant here are:

- (i) that subjection to them is nonvoluntary,
- (ii) that they are treated as having final authority, and

21 Kolodny’s analysis is very similar to Anderson’s analysis in terms of hierarchies of authority, esteem, and consideration. See Anderson, “Equality,” 43–45; and Anderson, *Private Government*, 3–4.

22 Kolodny, “Rule over None II,” 298.

23 Kolodny, “Rule over None II,” 298.

24 Although, as Kolodny notes, we should not be too sanguine about the workplace and other contexts not being sources of social inequality.

(iii) that they involve the use of force.²⁵

According to Kolodny, then, insofar as these are features of the political context, inequalities in power and *de facto* authority will suffice for relations of social inequality in that context.²⁶ These features plausibly explain why inequalities in power and authority are problematic in the political context but not in other contexts.

With these claims in place, we are now in a position to connect democracy and political equality to social equality: given features (i)–(iii), if there are asymmetries of political power and authority, then there is at least some failure to achieve the ideal of social equality.²⁷ So, the thought is that once we bear in mind these features of the political context, we will see that asymmetries of political power and authority (in the absence of certain dispositions) suffice for, or constitute, a form of social inequality. Under such conditions, those with greater power relate to those with less as superiors to inferiors. Insofar as we have an interest in standing in relations of social equality to those we share ongoing social relations with, we will have at least a *pro tanto* reason to support democracy over alternative forms of government since only democracy, by definition, involves equalities in (opportunities for) relative power and *de facto* authority.

2. THE CONSTITUTIVE DISTINCTNESS OF DEMOCRACY AND SOCIAL EQUALITY

While I find much to agree with in the egalitarian argument, I believe that the relation between democracy and social equality is not as tight as many egalitarians think it is. In this section I will begin by presenting an example that, I will argue, shows that mere inequalities in power and *de facto* authority do not suffice, even in the political context, for social inequality. This example tells directly against the simple view of the relation between democracy and social equality mentioned in the previous section. It does not, however, tell directly against Kolodny's version of the argument, but it does set the scene for an examination of Kolodny's more complex view. On the basis of an additional

25 Kolodny, "Rule over None 1," 226.

26 These features play a role in Kolodny's questioning, in his response to Anderson's arguments in *Private Government*, whether hierarchies in the workplace are as bad as they seem. See Kolodny, "Help Wanted," 107.

27 Daniel Viehoff understands Kolodny's argument in a similar way. Of arguments like Kolodny's, Viehoff writes: "Inequality in power is (unless qualified in certain quite specific ways) itself constitutive of social hierarchy, rather than being merely a causal antecedent of certain hierarchical social relations" ("Power and Equality," 11). The contrast here is between merely contingent and causal relations as opposed to necessary and constitutive relations.

series of cases, I will argue against the more complex view. I will argue that inequalities of power and *de facto* authority, together with the absence of certain dispositions, do not suffice for social inequality.

2.1. *Why Mere Inequalities in Power and De Facto Authority Do Not Suffice*

Consider a futuristic society that conforms to Kolodny's conception of a social democratic utopia: all members of the society have an equal opportunity to influence political decisions, and all relate to each other as social equals—no one is “above” or “below” anybody else. Using the wonders of modern technology, this society is able to ensure that voting on a range of issues is as easy as going to the local store. Indeed, it does this by making digital voting machines widely available throughout the society. The voting machines work perfectly, and everybody's vote is registered exactly once.

Now consider an almost identical society, which differs from Kolodny's social democratic utopia in only one respect: unbeknownst to everyone in the society, there is a serious fault with the voting machines. They register the votes of half of the citizens (those with names beginning with the letters *A* through *M*, say) correctly, but they register the votes of the other half of the citizens (those with names beginning with other letters, *N* through *Z*, say) twice. A consequence of this, let us suppose, is that the outcomes of votes in this second society never differ from those in the first hypothetical society. Because there is nothing suspicious about the outcomes of the decision-making process, nobody ever stops to wonder whether there is a fault with the voting machines, and nobody ever comes to know that the votes of some are counted twice.

It is fair to say that those with names beginning with the letters *N* through *Z* in the second society have greater power and *de facto* authority than those with names beginning with the letters *A* through *M*. To use Kolodny's vocabulary and image, some have greater *contributory influence*, where their greater influence can be modeled in terms of their applying a vector of force that has a greater magnitude than the vector of force applied by others in determining the result of the vote.²⁸ So, the two societies certainly differ with respect to the distribution of power and *de facto* authority. Now, the crucial question for us is this: Do these societies *also* differ with respect to how their members relate to each other? Do some relate to others as superiors to inferiors? By hypothesis, members of the first society relate to each other as equals. Do members of the second society relate to each other as equals, or do some relate to others as superiors to inferiors? It seems to me that there is *no difference* between the two societies with respect to *social equality*: even though some have more power

28 Kolodny, “Rule over None 1,” 200.

and *de facto* authority than others in the second society, they do not thereby relate to those others as superiors to inferiors.

This kind of example is a problem for the simple view of the relation between democracy and social equality but not for Kolodny's more complex view. The case involves the features that Kolodny claims make inequalities in power and *de facto* authority problematic: subjection is involuntary, the authority is final, and force is involved. It involves, after all, political decisions. Yet, even with these conditions in place, mere inequalities in political power and *de facto* authority do not suffice for social inequality. We seem to have an example on our hands where there are inequalities in political power and *de facto* authority and yet there is no social inequality. Political equality is not necessary for full social equality. The example does not, however, immediately tell against Kolodny's more complex view since, as we saw, it is important for Kolodny that inequalities in power and *de facto* authority be accompanied by the absence of particular dispositions. For Kolodny, social inequality involves some having greater relative power "while not being resolutely disposed to refrain from exercising that greater power as something to which those others are entitled."²⁹ There is nothing in our example to suggest that those who have the greater power are not resolutely disposed to refrain from exercising that greater power as something to which the others are entitled. Perhaps they are. (This is a point I will return to below.)

Before moving on to raise problems for Kolodny's more complex view, I want to consider whether the conclusion I have drawn from the example can be resisted and whether the simple view of the relation between democracy and social equality can be defended. How might my conclusion be resisted? Suppose someone disagrees with my verdict that there is no difference in how people relate to each other between the cases in the relevant sense. How might they do so? Well, it should be immediately conceded that the members of the second society do not relate to each other in exactly the same way as members of the first society do in a very broad sense of "relate to each other": even though they do not know it, some members of the second society relate to others as those with more power and *de facto* authority. The question for us, however, is whether *this* is a way of relating as superior to inferior or a way of relating as equals. And here it might be thought that this just *is* a way of relating as superior to inferior since having more power and *de facto* authority is characteristic of superiority. But this borders on question begging. There may be a sense in which some relating to others as those with more power and *de facto*

29 Kolodny, "Rule over None II," 295.

authority is a way of relating as superior to inferior. But this itself does not seem to be incompatible with them relating as equals in the relevant sense.

It would be disappointing if our disagreement here amounted to nothing more than a verbal disagreement about the application of the expression “social equality.” I think that we can make progress, however, by considering independently each of the characteristic features of social inequality identified by Kolodny. It should be very clear that in the examples provided, there are no asymmetries of consideration in the relevant sense. In terms of consideration, members of the second society certainly relate to each other as equals. The fact that asymmetries in consideration are a large component of what we object to when we object to the idea of relating as superior and inferior explains why we do not find the second society particularly problematic. Now let us ask whether there is anything particularly problematic about the asymmetries of power and *de facto* authority in the second society. On the face of it, there does not seem to be anything particularly problematic here.³⁰ Insofar as relations of social inequality are meant to be problematic, this apparent lack of a problem would suffice to show that there are no relations of social inequality here. Alternatively, we might concede that there does seem to be something problematic here, but then we might attempt to explain away the appearance. We might argue that our concern with asymmetries of power and *de facto* authority stems from our concern to avoid asymmetries in consideration. It is primarily because greater power and authority can be used to *extract* differences in consideration that we object to it. But for it to be so used, those who have it need to know that they have it, and those who do not need to know that they do not. Since those with more power and authority do not know that they have it in the second society, we correspondingly find the asymmetry either not problematic at all or far less problematic than in cases where there is a risk of the asymmetries being used to extract differences in consideration.

What the example seems to support, then, is what I will call the *extraction* theory of the bad of inequalities in power and *de facto* authority: these

30 I am not denying that asymmetries in power and authority are never problematic. Of course, there are paradigm examples of problematic power and authority, such as that of the kidnapper and their victim. In such cases, however, the asymmetry seems to be problematic because it allows those with greater power to interfere with the basic freedoms of their victim and to invade their body and property. I do not think that we are able to extrapolate from such examples to the case of problematic inequalities in power at the political level. If the state were to interfere in a person’s basic freedoms, or to invade their body and property, that would be problematic. It is the nature of state-subject relations that states have great power over their subjects and that there is a danger that this power might be used illegitimately. It seems to me that it is not some having greater power over the state’s decision-making in such cases that is problematic but rather the state’s interference.

inequalities are problematic not because, in certain circumstances, they constitute social inequalities but because, in certain circumstances, they can be *used* to *extract* greater consideration for those with greater power. Michael Walzer suggests a view like this when he writes:

It is not the fact that there are rich and poor that generates egalitarian politics but the fact that the rich “grind the faces of the poor,” impose their poverty upon them, command their deferential behavior. Similarly, it’s not the existence of aristocrats and commoners or of office holders and ordinary citizens (and certainly not the existence of different races or sexes) that produces the popular demand for the abolition of social and political difference; it’s what aristocrats do to commoners, what office holders do to ordinary citizens, what people with power do to those without it.³¹

The extraction theory can be contrasted with *constitutive* theories on which the bad of inequalities in power and *de facto* authority, in certain circumstances, *constitutes* or *suffices* for social inequalities. The example of the faulty voting machines seems to directly support the extraction theory over a simple constitutive theory of the bad of inequalities in power and *de facto* authority. Insofar as we find these inequalities problematic, it is because they can be *used* to extract greater consideration for those with greater power. However, since the example does not tell against Kolodny’s more complex version of the constitutive theory, it does not support the extraction theory over Kolodny’s theory. I will consider Kolodny’s theory in more detail in the next section. But before doing so, I want to fill out the extraction theory a little more in order to have something concrete to compare Kolodny’s theory with.

To fill out the extraction theory a little more, it might be helpful to return to Kolodny’s statement of the third characteristic feature of cases of social inequality. Kolodny’s original statement of the consideration condition is as follows:

- (iii) Some having attributes (e.g., race, lineage, wealth, perceived divine favor) that generally attract greater *consideration* than the corresponding attributes of others.

According to the extraction theory, one of the relevant attributes here might be perceived greater power. Some having perceived greater power might generally attract greater *consideration* than the corresponding attributes of others. The extraction theory effectively subsumes the inequalities in power condition

31 Walzer, *Spheres of Justice*, xii–xiii.

under the inequalities in consideration condition.³² Something like this view is suggested in the following passage from Miller:

Social equality is a matter of how people regard one another and how they conduct their social relations. It does not require that people be equal in power, prestige, or wealth. . . . What matters is how such differences are regarded, and in particular whether they serve to construct a hierarchy in which *A* can unequivocally be ranked as *B*'s superior.³³

For Miller, then, social equality is “a matter of how people regard one another and how they conduct their social relations”—that is, it is a matter of consideration. Inequalities in power, prestige, or wealth, according to Miller, are not in themselves problematic from the point of view of social equality. Rather, the issue is how these differences are regarded. Do those with perceived greater power, prestige, or wealth generally attract greater consideration? Do those with perceived greater power, prestige, or wealth *use* this perception to *extract* greater consideration? According to the extraction theory, differences in consideration will be largely due to those with greater power using that power to *extract* greater consideration from those with less power. It may well be, in addition, that those with greater power simply attract greater consideration without using their power to extract such consideration. Even in such cases it may be possible to speak of greater consideration being *extracted* from those with less power. I will say something about the difference between merely attracting greater consideration and extracting that consideration from others in a moment.

To fully fill out the extraction theory, more will need to be said about what consideration is such that some having attributes that generally attract greater consideration than the corresponding attributes of others constitutes a form of social equality. Following Kolodny, we can get an initial fix on the notion of consideration by thinking of those responses that social superiors, as social superiors, characteristically attract.³⁴ Such responses would include showing certain forms of deferential respect, fawning and toadying, bowing and scraping, other forms of deference, and efforts to ingratiate or curry favor.³⁵ We tend to recognize responses that social superiors, as social superiors, characteristically attract when we see them. It is very difficult to say, however, exactly what characterizes this class of responses. They are all inherently ways of relating as inferior to superior. They all involve treating someone as superior, even if those

32 And it focuses on *perceived* inequalities in power over *actual* inequalities in power.

33 Miller, *Principles of Social Justice*, 239.

34 Kolodny, “Rule over None II,” 297.

35 Walzer, *Spheres of Justice*, xiii; Kolodny, “Rule over None II,” 297.

who respond in these ways do not believe that the other is somehow above or superior. But what is it about these responses that *makes* them inherently ways of relating as inferior to superior? How do they differ from other forms of positive responses that are not inherently ways of relating as inferior to superior? As Kolodny notes, the class of responses that constitute consideration is partly characterized by the fact that although the basis of these responses is generally some narrow accidental feature of the person, the response is generally focused on the person themselves and their interests and claims.³⁶ Moreover, the responses are practical matters that deal with the person and claims of the person and have an agent-neutral character—the relevant feature is seen as calling for the same response from everyone.³⁷ However, as Kolodny notes, these conditions are not sufficient to characterize the relevant notion of consideration, and a full account remains elusive.³⁸ My own view is that relating as equals in our social interactions is largely a matter of *experiencing ourselves as being equals* in our social interactions. Some ways of responding to others involve an *experience* of subordination, or inferiority. Here I am taking seriously Walzer's claim that "the experience of subordination—of personal subordination, above all—lies behind the vision of equality."³⁹ This experience is not a response to a perception of superiority in another; rather, it belongs to the form of response itself. It is this experience of inferiority, I think, that justifies talk of the *extraction* of greater consideration. The consideration is felt as being extracted from the inferior and granted to the superior.⁴⁰ We might think of a class of responses that constitute consideration as practical responses to the whole person that are based on perceived accidental features of the person and that have an agent-neutral character and are accompanied by an experience of inferiority on the part of those whose responses they are. What I have offered here is an initial sketch of what a theory of consideration might look like and how it might be developed. Given my aims in this paper, I will not attempt to develop the theory in any more detail here. For my purposes,

36 Kolodny, "Rule over None II," 298.

37 Kolodny, "Rule over None II," 298.

38 Kolodny remarks that while a freestanding account of social inequality would need an analysis of consideration, he does not need one since his argument for democracy will focus on the inequality in power condition and its relation to democracy ("Rule over None II," 298).

39 Walzer, *Spheres of Justice*, xiii.

40 This view might be thought to run into trouble with false consciousness. But I think that such trouble can be avoided by noting that even when one believes that greater consideration is due to one's superiors, one will experience a sense of inferiority in responding to them in ways that constitute giving them greater consideration.

it is enough to have a clear view of the alternative approach that I want to contrast Kolodny's theory with.

To sum up the argument to this point, I have argued that the example of the faulty voting machines poses a problem for the simple view of the relation between democracy and social equality. Inequalities in power and *de facto* authority appear to be compatible with full social equality even in the political context. I have just suggested that any lingering sense that these inequalities are problematic may be explained on the hypothesis that we think that such inequalities are problematic because when those who have greater power know that they do, they can use their greater authority and power to extract greater consideration from others, thus undermining social equality. Now, someone like Kolodny could well agree with the conclusions reached so far. He could agree that mere inequalities in power and *de facto* authority do not themselves suffice for social inequality. And he could agree that such inequalities are often problematic because they can be used to extract differences in consideration. However, Kolodny will want to say something stronger also. He will want to claim that such inequalities, together with the absence of certain dispositions, *suffice* for a kind of social inequality. If so, then there would be a reason, over and above the fact that such inequalities can be used to extract greater consideration, for objecting to inequalities in power and authority. And we would have a less empirical and contingent reason in favor of democracy over the alternatives.

2.2. *On the Absence of Dispositions to Refrain from Exercising Power*

What does Kolodny mean when he says that social inequality involves some having greater relative power over others "while not being resolutely disposed to refrain from exercising that greater power as something to which those others are entitled"? We can start to get a feel for what he means by reflecting on why, in the example of the faulty voting machines, even though half of the people in the society *do* exercise their greater power and *de facto* authority, they may nonetheless possess the disposition Kolodny is referring to. In the example of the faulty voting machines, half of the people in the society exercise their greater power and *de facto* authority *unknowingly*. They are disposed to exercise their greater power *unknowingly*. But this disposition is perfectly compatible with their being resolutely disposed to refrain from exercising their power *knowingly*. It may be that if those with greater power discovered that they had it, they would immediately stop exercising it and seek to have equality of power restored.

This explains why the example of the faulty voting machines does not pose an immediate problem for Kolodny's view. For all that we have said, members of the society might be resolutely disposed to refrain from exercising their greater power as something to which others are entitled. Now, at this stage,

we might wonder what would happen if we simply stipulated that those members of the society are *not* so disposed. Consider, then, the first society, the society without faulty voting machines. Suppose each member of that society *would* happily exercise greater power over others if they could get it. But as things stand, there is no means for them to get more power, and they are not particularly interested in trying to get it. While this may reflect badly on their character, it is doubtful that merely not being resolutely disposed to refrain from exercising greater power over others would suffice for a form of social inequality. Suppose that we carry this condition over to the society with faulty voting machines, again supposing that members of the society are not aware of the fault. Do we not now have an example where there are inequalities of power and *de facto* authority, together with the absence of a disposition to refrain from exercising such greater power and authority? Well, yes and no. There is a sense in which those with greater power lack a resolute disposition to refrain from exercising it; they are just not in a position to exercise the power knowingly since they do not know that they have it. But it would be uncharitable to understand Kolodny's position this way. It would be more charitable to understand Kolodny as holding that those who have the greater power are such that (i) they are in a position to exercise it knowingly and (ii) they are not resolutely disposed to refrain from exercising their greater power as something to which others are entitled. This gives a deeper reason for thinking that examples like the faulty voting machines, which turn on the absence of knowledge of greater power, do not pose an immediate problem for views like Kolodny's. For Kolodny, greater power is problematic, in part, when those who have it are in a position to exercise it knowingly and yet are not resolutely disposed to refrain from doing so for the right reasons.

While Kolodny's theory is not immediately threatened by the examples we have considered so far, I think that there are related examples that *do* pose an immediate problem for it. Consider a variation on the faulty voting machines example where the fault is due to a devious programmer. The programmer studied political philosophy for many years and, being unable to find an academic job, turned to programming and became involved in programming software for voting machines. After a few years of running his social experiment with the faulty voting machines, he decides to change things up a bit. He secretly contacts those whose votes are counted twice, letting them in on this fact. Since he explains his role, those who he contacts have good reason to take him at his word. Now, the programmer is careful to ensure that each of those he contacts does not know who else's votes are counted twice. And those whose votes are counted twice have good reason to think that if they tried to discover who else had greater power, they would come undone, and the whole system would be

undermined. Suppose, as in the variation on the example we were discussing a moment ago, that members of this society are not resolutely disposed to refrain from exercising their greater power knowingly. Indeed, the programmer may have explained to them that since those with greater power have basically been randomly selected, their exercising their greater power makes no difference to the outcome of the voting procedure. In light of this, those with greater power may think that it does not really matter that they have and exercise greater power. They exercise it more from indifference than malice.

We now have an example where all of Kolodny's conditions for social inequality are met: subjection is involuntary, the authority is final, force is involved, and some have greater power and *de facto* authority while not being resolutely disposed to refrain from exercising that power and authority as something to which others are entitled. Yet, intuitively, there is no more social inequality in this society than in our original society. There is no relevant difference in how members of this society relate to each other. There is no reason to think that those who choose to exercise their greater power in the circumstances hold a malevolent attitude to other members of their society. Those who know they have greater power and *de facto* authority are in fact rather lonely in their knowledge since they know that as soon as they seek to share their knowledge, the whole system will unravel. It is their lonely secret. Moreover, in light of the fact that they also know that their having this greater power makes no difference to the outcomes, it would be hard for them to develop a sense of superiority over others. Having greater power is nothing to them.

If it is agreed that there is no more social inequality in this society than in our original society—the one in which the voting machines work as they should—then we have a counterexample to Kolodny's position *as stated*. There are, however, two paths of response open at this stage: a proponent of the constitutive theory might resist the verdict about the case, or they may argue that a small adjustment can be made to the theory in order to avoid this kind of example while still retaining the spirit of the theory. I will examine each kind of response in turn.

There are two ways that a proponent might try to resist my verdict about the case. They might just bluntly state an opposed intuition about the case. Or they might argue that the case is sufficiently similar to cases where we judge that there is a kind of problematic social inequality involved. Suppose someone bluntly states an opposed intuition. Then we can consider similar arguments to those we considered in the previous section. We might ask, What is meant to be so problematic about the circumstances in this society? There may be a lingering sense of social inequality here. But again, we should consider what is meant to be problematic about this combination of greater power and the

absence of certain dispositions. The example under consideration is designed not only as an example where Kolodny's conditions are met but also as an example where we have ensured that those with greater power and authority are not in a position to extract greater consideration *even though* they know they have greater power and authority. By hypothesis, if they were to try to extract greater consideration, the whole system would come undone. So, they cannot use their greater power to extract greater consideration. Even if we think that merely having greater power while lacking the relevant disposition is problematic, it is nowhere near as problematic as when that greater power can be used to extract greater consideration. So, again, we should consider a deflationary explanation of our intuitions about inequalities in power in the absence of certain dispositions amounting to social inequality. Indeed, once we are considering such dispositions, we should be even more tempted by an explanation in terms of the extraction theory. For it is plausible to think that those who would not be resolutely disposed to refrain from exercising their greater power, in ordinary cases, would also not be disposed to refrain from extracting greater consideration on the basis of their having greater power. There is, however, a risk that this line of reasoning may simply lead to a clash of intuitions about the case. So let us now examine how a proponent of the constitutive theory might offer an argument to resist the verdict about this case.

A proponent of the constitutive theory might argue that the case at hand is sufficiently similar to a kind of case that is clearly problematic and that clearly involves social inequalities. They may conclude from this that the case at hand involves social inequalities, but perhaps to a lesser degree. The kind of case the proponent of the constitutive theory will appeal to will be a case involving rule by a secret society or the Illuminati, for instance. To bring this case as close as possible to the case at hand, we might imagine that the Illuminati has hired our programmer to introduce the bias into the voting system and to have the votes of members of the Illuminati counted twice. And, of course, every other member of the society would be left in the dark. Now, we certainly would find such an example problematic. There is something objectionable about secret rule by the Illuminati. And the example has similarities to the case at hand. Those with less power are not aware that those with greater power have greater power, and if their power were to be discovered, the whole system would collapse. Given that we find this example so problematic, why should we not find the original example of the devious programmer problematic? Well, there is an important disanalogy between the example of the Illuminati and the original example. Since the members of the Illuminati are mutually aware of their having greater power and can communicate their intentions with nods and winks without being noticed, they can be said, collectively, to exert a kind of

intentional control over the decision-making process that the individuals with more power, taken collectively, could not be said to exert in the case at hand. I suspect that this is what drives our intuitions about the Illuminati example. And since it is missing in the original example of the devious programmer, our intuitions about the Illuminati example should not be taken to bear on the original example.

An interesting question remains, however, as to whether we think that the Illuminati example is problematic because it involves relations of social inequality or whether we think it is problematic for some other reason. Do members of the Illuminati relate to other members of the society as *social* superiors? My sense is that they do not. Here is why: it does not seem to matter for our verdict whether members of the Illuminati are members of the society they control or whether they are members of another society, ruling the original society remotely. So the issue does not seem to be one about *how* members of a society relate to each other at all. Now, it might be said that, nonetheless, when those with greater power who are members of a collective that controls the decision-making process *are* members of the same society they control, they are *thereby* socially superior. This is fine, but it amounts to a stipulative use of “socially superior,” and what is problematic about *this* instance of social superiority would be shared with cases of foreign rule and is not explained in terms of a failure to relate as equals. In any case, since those with greater power in the original example are not members of a collective that exerts a kind of *intentional control* over the decision-making process, this line of reasoning does not provide an argument against our verdict about this example.

The second way a proponent of the constitutive theory might respond to the example of the devious programmer is by arguing that a small adjustment can be made to the theory in order to avoid this kind of example while still retaining the spirit of the theory. The proponent of the constitutive theory might think that the lesson of the first example of the faulty voting machines is that in order to have greater power or *de facto* authority over others in the relevant sense (i.e., in the sense that is constitutive of social inequality), it is not merely enough that one have *greater contributory influence*. It requires that one *know* that one does. In response to the example at hand, they might argue, similarly, that in order to have greater power or *de facto* authority over others, it is not enough to *know* that one has greater contributory influence; those who have lesser contributory influence must *know* that you do. In other words, perhaps the kind of inequalities in power that are constitutive of social equality are inequalities of power that are common knowledge. At the very least, it might be thought that in order for inequalities in power to be constitutive of social inequality, those with less power must have an inkling that some have more power than they do.

So why not amend Kolodny's statement of the conditions that constitute social inequality as follows: some having greater relative *power* (whether formal or legal or otherwise) over others, while not being resolutely disposed to refrain from exercising that greater power as something to which others are entitled, where those with lesser relative power have an inkling that those with greater relative power have greater relative power.

Would this amendment not be largely in the spirit of the constitutive theory and save it from the counterexample? Is it not just obvious that this is what a proponent of the constitutive theory should say? My view is that it will not do simply to amend the theory in this way. What the constitutive theorist needs is that inequalities in power themselves be at least *in part* constitutive of social equality. The constitutive theorist can amend the theory by claiming that it is inequalities of power *together with* an inkling on the part of those at the losing end of the inequalities that constitute social inequality. But this opens them to two objections. First, why should adding knowledge or suspicion on the part of those at the losing end to inequalities in power suddenly *amount* to a failure of social equality? What does mere suspicion of inequalities of power add to inequalities of power such that together they suddenly *constitute* social inequalities? Would not a better explanation be that the suspicion of inequalities of power is *causally* relevant to social inequality? Would not those who suspect that they have less power than others fawn over those with more power and give them greater consideration? Second, once this possibility is noticed, there is a threat that inequalities of power themselves will drop out of the picture as irrelevant. Either it is the suspicion of greater power that constitutes the social equality or it is the suspicion of greater power that is causally relevant to social equality. Either way, there is no work for *actual* inequalities in power to do in constituting social inequality. I will now present an example that, I think, shows that inequalities in power do, in fact, drop out of the picture as irrelevant in either case.

2.3. *Why Inequalities of Power and De Facto Authority Add Nothing*

Consider a final variation on the faulty voting machine example. This time, begin with a society where, by design, the system is meant to be a system of plural voting. In this society it is common knowledge that the votes of some are counted twice while the votes of others are counted only once. Moreover, and predictably, those with more power use their power to extract greater consideration, so there is significant social inequality supported by the system. Those with more power relate to those with less as superiors to inferiors. But now here is the twist. Our devious programmer has been at it again, and while it is widely believed that the votes of some count for more than the votes of others, this is not in fact the case, and everyone's vote is counted only once. So

one means by which those with perceived greater power might have actually had greater power has been closed off. This is not to say, however, that in this example everybody has equal power. It may well be that those with perceived greater power nonetheless *do* have slightly greater informal power since they will be able to influence the decision-making of others through their perceived greater power.

Let us now ask if things would be *worse* from the point of view of social equality if the programmer had done his job correctly and programmed the plural voting scheme. That is, let us ask if things would be worse if there were greater formal inequality of power and authority in the society. If the programmer had done his job correctly, then in addition to the social inequality due to the extraction of greater consideration and the purported inequalities due to whatever greater power the perception of greater power gives to those who are perceived to have it, there would be *additional* social inequality constituted by some having, through the formal voting system, greater power and *de facto* authority while not being resolutely disposed to refrain from knowingly exercising it. But, on the face of it, things would *not* be worse from the point of view of social equality if this were the case. All the damage is already done by the widespread *belief* that some have more power and authority. No further damage is constitutively done by there *actually* being additional asymmetries in power and authority accompanied by the absence of certain dispositions. On the plausible assumption that this is not simply a case of overdetermination, we should conclude that inequalities of power and *de facto* authority do not suffice in the circumstances for additional social inequality.

This case confirms the suspicion raised at the end of the previous section that once we add knowledge or belief to the picture, the asymmetries in power due to actual differences in contributory influence drop out of the picture as irrelevant. And this should lead us to see the earlier example of the devious programmer in a new light. There we introduced the hypothesis that it was really what those with greater power could do with power that we objected to and that led to social inequality. Now we can see that they need not actually have greater power. There only needs to be a mutual perception that they have greater power.

This completes my case against both the simple view of the relation between democracy and social equality and Kolodny's more complex view of this relation. Neither mere inequalities in power and *de facto* authority nor inequalities in the absence of certain dispositions suffice in the political context for social inequality. When those lacking a disposition to refrain from exercising greater power have greater power over others but are not in a position to extract greater consideration, there is no threat to social equality, and merely having greater power while lacking the relevant disposition does not suffice for social equality.

Moreover, when some are in a position to extract greater consideration on the ground of their supposedly having greater power and authority, their actually having greater power and authority, and their lacking the disposition to refrain from exercising it, does nothing constitutively to further undermine social equality.

At this stage it might be worth briefly returning to consider Kolodny's motivating example, the society of ascetic warriors. Why might we have been initially inclined to agree that in such a society there was a problematic form of social inequality? Well, I suspect that it is partly because the example is underdescribed. Kolodny is eager to describe the case as one in which the requirements of distributive justice are met. But we know little else about how members of the society respond to rule by the ascetic warriors. Perhaps we would not judge the example to be problematic if it came closer to our examples of the faulty voting machines, and there was some factor ensuring that the warriors would not attract greater consideration in virtue of their perceived greater power. Perhaps we judge the example to be problematic only because we assume that their having perceived greater power would attract greater consideration. We cannot conclude immediately on the basis of considering this example that it is the inequality in power and *de facto* authority that constitutes the problematic social inequality. We need examples like those we have been considering that attempt to tease apart inequalities in power and *de facto* authority from inequalities in consideration. I have argued that when we consider such examples, we see that it is ultimately the inequalities in consideration, inequalities that can be extracted by those with perceived greater power from those with lesser power, that constitute the problematic forms of social inequality.

3. A COMPARISON

In this section I compare the argument given above with a related argument offered by Richard Arneson in the context of discussing Philip Pettit's work on nondomination.

The argument I have just given is structurally similar to an argument Richard Arneson briefly sketches in a discussion of republican freedom, or freedom as nondomination (in the broader context of a discussion of social equality). As Arneson notes, a central theme in Pettit's work is that freedom as nondomination is instrumentally valuable as a means to avoiding certain relations of social inequality.⁴¹ For Pettit, the connection between domination and social inequality is mediated by common knowledge. Here is the crucial passage from Pettit:

41 Arneson, "Democratic Equality and Relating as Equals," 46.

Domination is generally going to involve the awareness of control on the part of the powerful, the awareness of vulnerability on the part of the powerless, and the mutual awareness—indeed, the common awareness among all the parties to the relationship—of this consciousness on each side. The powerless are not going to be able to look the powerful in the eye, conscious as each will be—and conscious as each will be of the other’s consciousness—of this asymmetry. Both will share an awareness that the powerless can do nothing except by the leave of the powerful: that the powerless are at the mercy of the powerful and not on equal terms. The master-slave scenario will materialize, and the asymmetry between the two sides will be a communicative as well as an objective reality.⁴²

For Pettit, then, domination is a means to social inequality.⁴³ When it is common knowledge that there are relations of domination, these relations can be used to extract differences in consideration, in social status. Having drawn attention to this aspect of Pettit’s view, Arneson then writes: “This comment on the badness of dominating power raises the question: suppose the [conditions for domination] obtain but nobody knows this is so, or nobody knows but the dominant party, and domination is never exercised, so none of the envisaged bad consequences occur.”⁴⁴ The first part of what Arneson is asking us to suppose is similar to what we supposed in the initial version of the faulty voting machines example: “suppose nobody knows this is so.” The question Arneson then asks is whether, on this supposition, relations of domination would be noninstrumentally bad. He reports the intuition that they would not be, and in light of the foregoing discussion of the faulty voting machines example, we can agree. Arneson draws the following moral: “Pettit’s discussion calls attention to the fact that one might object to inequality of power without prizing equality of power *per se*.”⁴⁵

In the previous section, I argued that Kolodny’s position is not directly susceptible to this style of argument. For Kolodny, asymmetries in power and *de*

42 Pettit, *Republicanism*, 60–61.

43 Pettit returns to the theme later in his discussion. See Pettit, *Republicanism*, 71–72. Speaking later of the instrumental relation to social equality mediated by a “shared awareness of the asymmetry of power” (88), Pettit writes: “The enjoyment of freedom as non-domination goes with the possibility of their seeing themselves as non-vulnerable in that way and as possessed of a comparable social standing with the other. They can look the other in the eye; they do not have to bow and scrape” (87). Also: “To be able to live without subordination to others,” “these are great and palpable goods and they make a powerful case for the instrumental attractions of freedom as non-domination” (90).

44 Arneson, “Democratic Equality and Relating as Equals,” 46.

45 Arneson, “Democratic Equality and Relating as Equals,” 47.

facto authority are not problematic in and of themselves. They are problematic only in conjunction with the absence of certain dispositions. We have seen, however, that this style of example can be extended to target Kolodny's position, and we ended up with a view very similar to Pettit's here. We concluded that asymmetries in power and *de facto* authority are problematic primarily because they can be used to extract additional consideration for those with perceived greater power. When it is common knowledge that they obtain, asymmetries of power and *de facto* authority can be *used* as means to social inequality—a point Pettit makes well in his discussion of the instrumental value of nondomination.

It is striking that Kolodny does not comment on this aspect of Pettit's view when he compares his own approach to Pettit's. In his discussion of Pettit on freedom as nondomination, Kolodny suggests—rightly in my view—that our concern with nondomination may be not so much with being under an arbitrary and alien will itself but with being “on the losing end of an asymmetry of power with another person.”⁴⁶ Given the passage I quoted above, it may be plausible to reinterpret Pettit as being concerned more with asymmetries of power than with being under an arbitrary and alien will.⁴⁷ However, reinterpreting Pettit's concerns in this way would see asymmetries in power and *de facto* authority as potential *means* to social inequality, and not as sufficing for or constituting social inequality in the political context. After all, Pettit insists that inequalities in power would be problematic even in the presence of a robust disposition not to knowingly exercise them. From our point of view, this looks plausible since those with greater power could still use that power to extract greater consideration *even if* they were in fact resolutely disposed not to knowingly exercise that power.

4. CONCLUSION

Where does our discussion leave the egalitarian argument for democracy? Well, in my view, the foregoing does not have any significant implication for the conclusion of the argument—namely, that we have reason to prefer democracy over other forms of government. Nor does it give us any reason to think that there could be no good argument from considerations about social equality for democracy. It rather has implications for the *kind* of argument we might hope for here.

46 Kolodny, “Being under the Power of Others,” 106.

47 Kolodny remarks on the appeal to social equality in Pettit's work in Kolodny, “Being under the Power of Others,” 107–8. See Pettit, *On the People's Terms*, 83–85, for the “eyeball” test.

For Kolodny, since democracy is necessary, in a strong sense, for social equality, insofar as we have an interest in social equality, this interest gives us a reason to prefer democracy over alternatives to democracy. There is no need for an empirical argument to the effect that democracy is, contingently, as things are here and now, the only available means to avoiding social inequality. It is enough for social inequality, given the nature of the political context, that there be differences in power and *de facto* authority in the absence of certain dispositions. We can know from the armchair that inequalities in political power and *de facto* authority, in certain circumstances, suffice for and partly constitute social inequality.

The examples above suggest that the case for democracy on the basis of considerations of social inequality will depend, in large part, on the empirical premise that inequalities in power and *de facto* authority will be used to extract greater consideration for those with greater power. Now, in my view, this empirical premise is very plausible. But to establish it, we would need to do more than just reflect on hypothetical examples. We would need to investigate the conditions that allow those with greater power to extract greater consideration. It may well be that there are ways of arranging society that would allow some to have greater power and authority without thereby being in a position to extract greater consideration. Then our argument would not necessarily support democracy. But it is plausible to think that given the significance of political power and the importance of political decisions—something highlighted by Kolodny’s focus on finality, force, and involuntariness—the temptation to extract greater consideration would be too great to resist, and that unless we were not to relate to each other in society at all, greater consideration could easily be extracted from those with less power by those with more power.⁴⁸

University of Sydney
ryan.cox@sydney.edu.au

REFERENCES

- Anderson, Elizabeth. “Equality.” In *The Oxford Handbook of Political Philosophy*, edited by David Estlund, 40–57. Oxford: Oxford University Press, 2012.
- . *Private Government: How Employers Rule Our Lives (and Why We Don’t Talk about It)*. Princeton, NJ: Princeton University Press, 2017.

48 I would like to thank Matthew Hammerton, Luara Ferracioli, Daniel Te, and two anonymous referees for the journal for their helpful comments and feedback on this paper.

- . “What Is the Point of Equality?” *Ethics* 109, no. 2 (January 1999): 287–337.
- Arneson, Richard J. “Democratic Equality and Relating as Equals.” In “Justice and Equality,” edited by Colin MacLeod. *Canadian Journal of Philosophy Supplementary Volume* 36 (2010): 25–52.
- Benn, S. I., and R. S. Peters. *Social Principles and the Democratic State*. London: George Allen and Unwin, 1959.
- Christiano, Thomas. *The Constitution of Equality: Democratic Authority and Its Limits*. Oxford: Oxford University Press, 2010.
- Christiano, Thomas, and John Christman. *Contemporary Debates in Political Philosophy*. Malden, MA: Wiley-Blackwell, 2009.
- Crosland, C. A. R. *The Future of Socialism*. London: Cape, 1964.
- Dworkin, Ronald. “What Is Equality? Part 1: Equality of Welfare.” *Philosophy and Public Affairs* 10, no. 3 (Summer 1981): 185–246.
- Fourie, Carina, Fabian Schuppert, and Ivo Wallimann-Helmer, eds. *Social Equality: On What It Means to Be Equals*. New York: Oxford University Press, 2015.
- Kolodny, Niko. “Being under the Power of Others.” In *Republicanism and the Future of Democracy*, edited by Yiftah Elazar and Geneviève Rousselière, 94–114. Cambridge: Cambridge University Press, 2019.
- . “Help Wanted: Subordinates.” In *Private Government: How Employers Rule Our Lives (and Why We Don’t Talk about It)*, by Elizabeth Anderson, 99–107. Princeton, NJ: Princeton University Press, 2017.
- . “Rule over None I: What Justifies Democracy?” *Philosophy and Public Affairs* 42, no. 3 (December 2014): 195–229.
- . “Rule over None II: Social Equality and the Justification of Democracy.” *Philosophy and Public Affairs* 42, no. 4 (December 2014): 287–336.
- Miller, David. “Complex Equality.” In *Pluralism, Justice, and Equality*, edited by David Miller and Michael Walzer, 197–225. Oxford: Oxford University Press, 1995.
- . “Equality and Market Socialism.” In *Market Socialism: The Current Debate*, edited by Pranab K. Bardhan and John E. Roemer, 298–314. New York: Oxford University Press, 1993.
- . *Principles of Social Justice*. Cambridge, MA: Harvard University Press, 1999.
- Pettit, Philip. *On the People’s Terms: A Republican Theory and Model of Democracy*. Cambridge: Cambridge University Press, 2012.
- . *Republicanism: A Theory of Freedom and Government*. Oxford: Oxford University Press, 1997.
- Scheffler, Samuel. “The Practice of Equality.” In *Social Equality: On What It*

- Means to Be Equals*, edited by Carina Fourie, Fabian Schuppert, and Ivo Wallimann-Helmer, 21–44. New York: Oxford University Press, 2015.
- . “What Is Egalitarianism?” *Philosophy and Public Affairs* 31, no. 1 (Winter 2003): 5–39.
- Tawney, R. H. *Equality*. London: George Allen & Unwin, 1931.
- Viehoff, Daniel. “Power and Equality.” In *Oxford Studies in Political Philosophy*, vol. 5, edited by David Sobel, Peter Vallentyne, and Steven Wall, 3–38. New York: Oxford University Press, 2019.
- Wall, Steven. “Democracy and Equality.” *Philosophical Quarterly* 57, no. 228 (April 2007): 416–38.
- Walzer, Michael. *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books, 1983.
- Williams, Bernard. “The Idea of Equality.” In *Problems of the Self*, 230–49. Cambridge: Cambridge University Press, 1973.
- Wolff, Jonathan. “Equality and Hierarchy.” *Proceedings of the Aristotelian Society* 19, no. 1 (April 2019): 1–23.
- . “Fairness, Respect, and the Egalitarian Ethos.” *Philosophy and Public Affairs* 27, no. 2 (Spring 1998): 97–122.

WHAT GOES ON WHEN WE APOLOGIZE?

Christopher Bennett

APOLOGY is often said to play an important role in reconciliation. On a plausible interpretation of that claim, apology has this important role because the performance of an apology provides us with new practical reasons, reasons to change the way we relate to the wrongdoer. But if this is right, what kinds of reasons are they, and why is an apology necessary (and sufficient) to provide us with such reasons?

In this paper, I argue that our practice of giving and demanding apologies is underpinned by a belief that apologies make a difference to our normative situation: that once an apology has been given, the rights and responsibilities of the apologizer and others have been altered. However, if we ask what rationalizes that belief, two influential views in the literature on apology—which I call the *reassurance view* and the *performing deference view*—prove to be inadequate. One thing a theory of apology needs to explain is that the distinctive work of reconciliation carried out by apology involves a set of canonical actions through which one can change one's normative situation in characteristic ways. However, it is also a characteristic feature of apology that (at least in cases of serious wrongdoing) it effects this reconciliation only when sincere—that is, when it is an expression of the wrongdoer's remorseful recognition of the wrongness of what was done.¹ The reassurance view and the performing deference view fail to offer an explanation that is adequate to both of these features. In order to explain these features of apology, I suggest that we see apology as a power to change one's normative situation through the performance of canonical actions, but a power that is exercised expressively, or by an *expressive action*.²

For an apology to do its distinctive work, I will argue, it needs to be an expression of emotion that is *appropriate*. This means not only that the *emotion* should be appropriate but also that the emotion should find an appropriate

1 Govier and Verwoerd, "The Promise and Pitfalls of Apology."

2 In offering an account of apology as an expressive action, I will be drawing on an account of expressive action I have developed in a number of other recent papers. See Bennett, "Expressive Actions," "The Problem of Expressive Action," and "How and Why to Express the Emotions."

expressive *vehicle*. This, it will turn out, means that the expression of the relevant emotion has to conform to certain canonical features. Once we understand apology as an expressive action, we can see how, in favorable conditions at least, the canonical actions through which the apologizer's normative situation can be altered could also be those through which individuals can give authentic expression to their emotion. In this way, we can explain how it might make sense to see apologies as providing new practical reasons while at the same time expressing the wrongdoer's authentic remorse.

I

There are many things it is appropriate for a wrongdoer to do after wrongdoing. What I am interested in is understanding the normative work done specifically by apology. We can focus on this question by asking what is missing when an apology fails to be given. Imagine someone saying the following:

Missing Apology: I know that he's sorry for what he has done. And I am sure that he wouldn't do it again. I trust him and don't want to lose him as a friend. In fact, I think that despite everything that happened we *are* still good friends. I know he has changed since then, and he knows that I know it too. It just feels a bit strange that he has never apologized. It has just left everything a bit unresolved. Like there is unfinished business. And somehow I just can't feel entirely good about my relationship with him anymore because of it. I am sure he really does feel bad about what he did to me. But why won't he just come out with it?

The speaker in this example is insisting that *an apology* is what they are owed and that until an apology is forthcoming, matters are not settled. This is clearly not to say that only apology matters. The speaker is grateful that their friendship with the wrongdoer has survived. And they have an appreciative attitude to whatever the wrongdoer has done to rebuild trust between them. However, an apology has not been given, and this has left our speaker with the perhaps hard-to-pin-down feeling that something remains unfinished. In this section, I attempt to articulate what lies behind the speaker's feeling that there is unfinished business where an apology is not forthcoming.

To start with, let us get clear on what apologizing involves. I am interested in what we might think of as an unreserved apology.³ Looking at what is involved in apologizing unreservedly—where the apology is sincerely given and

3 See Davis, "On Apologies," on the "consummate apology." See also Smith, "The Categorical Apology."

unconditional—is relevant here because it can help us to see what elements are required for an apology to leave no business unfinished. In the following paragraph, I aim simply to summarize the features standardly given in accounts of such full, unreserved apologies.⁴

First of all, such apologizing normally involves speech addressed to the wronged party—an “I’m sorry.” However, as is often said, sorry is not good enough; so, second, this speech should involve an acknowledgment of responsible wrongdoing (“It was my fault”) and, third, a credible commitment to refrain from such acts in the future (“I can see that it was wrong, and I won’t do it again”). Fourth, for cases of nontrivial wrongdoing, apologizing unreservedly involves the wrongdoer showing that they are troubled by the wrong they have done to their victim; the apology is thus an expression of remorse. Apologies for nontrivial matters are undermined by the appearance that the apologizer does not feel remorseful for what they have done.⁵ Fifth, apology will also usually involve an offer of restitution and a commitment to make amends. Sixth, a person who is apologizing unreservedly does not do so stridently, confidently, unabashedly; rather their demeanor, posture, and gestures exhibit deference and humility. Seventh, a successful apology involves some credible attempt to make the extent of the remorse—the amends offered and the degree of humility—proportional to the perceived seriousness of the wrong. The wrongdoer thinking that an admission of responsibility and a commitment to refrain in the future are sufficient, or offering to make amends but only to a negligible degree, can reveal that they underestimate the seriousness of the wrong. An apology can thus misfire if it fails to include one or more of these seven elements (at least when it could reasonably have included them) or if what is offered fails to be proportionate to the seriousness of the wrong.

So, how can we start to articulate what lies behind the view expressed in Missing Apology? First of all, Missing Apology seems to imply that among appropriate responses to wrongdoing, it is reasonable to give and ask for apologies in particular. As we have noted, this is not to say that apology is the only appropriate response to wrongdoing. It is not even to say that it is the most

4 In addition to the papers already cited, see, e.g., Tavuchis, *Mea Culpa*; Joyce, “Apologizing”; Gill, “The Moral Functions of an Apology”; Lazare, *On Apology*; Smith, *I Was Wrong*; Bovens, “Apologies”; Bennett, *The Apology Ritual*; Martin, “Owning Up and Lowering Down”; Pettigrove and Collins, “Apologizing for Who I Am”; and Helmreich, “The Apologetic Stance.” In defending the possibility of vicarious and collective apologies, Andrew Cohen argues that apologies should be thought of in terms of characteristic functions rather than required features (see “Vicarious Apologies as Moral Repair”).

5 For some discussion of counterexamples that involve actions for which one says sorry but that were ultimately beneficial and thus not, it is claimed, cause for regret, see Barnum-Roberts, “Apologizing without Regret.”

important one. But it is to say that a wrongdoer's response to wrongdoing is incomplete if an apology is missing. It seems to be a corollary of this that the giving of an apology makes a moral difference that could not have been brought about by other means. However, if we want to say precisely what this moral difference is, we should note that Missing Apology also implies that while apology has a role in bringing about reconciliation after wrongdoing, it is neither necessary nor sufficient for actual resumption of good relations. Apology is not *necessary* for good relations since it is realistic to think that although the speaker is aware that something is missing, the pair in Missing Apology are getting along *okay*, with trust and goodwill and friendship. And apology is not *sufficient* for good relations because apology cannot bring about good relations in the absence of whatever additional trust-, friendship-, and goodwill-(re)building measures the wrongdoer in this case has undertaken. Apology cannot bring trust about magically, just by the uttering of "I am sorry; it's my fault; I won't do it again." Despite this, the speaker implies that apology is necessary and sufficient for *some* kind of closure after wrongdoing, since they think that some sort of closure is lacking because an apology has not been given.

I suggest that we look for what is distinctive in apology, not in its role in bringing about actual good relations, but rather in the way that it alters the normative situation that arises from wrongdoing. We can interpret the dissatisfaction expressed in Missing Apology as evincing an awareness that even though he may have done many other things that are appropriate, the wrongdoer has not brought about some alteration of the normative situation that apology can (and should) bring about. The speaker is in two minds because they recognize the good in the wrongdoer, and the trust that has been rebuilt, but nevertheless feel dissatisfied because there is an element of moral compromise in resuming relations in the absence of an apology, going forward as if everything were normal. Even though he has rebuilt confidence, the wrongdoer has not done that specific thing that would allow our speaker, in clear conscience and while doing full justice to the significance of the wrong, to resume normal relations with him. The speaker feels themselves to be in a situation of moral compromise because of the lack of an apology.

Now, in trying to make sense of Missing Apology, it is crucial to note that it is possible, as I have done, to specify *in advance* of any particular apology the elements that such an apology will need to involve in order to bring about the relevant kind of closure. As moral agents familiar with the normative expectations involved in giving and demanding apologies, we know in advance that apologies that are not addressed to the victim (when they easily could be), that do not involve a commitment to refrain in the future, that are not appropriately remorseful and deferential and proportionate, etc. will not cut the mustard.

This suggests an important explanandum for a theory of apology: that, whatever moral work is done by an apology, the elements required to make it normatively effective are relatively unchanging across contexts. Whereas rebuilding trust, or salvaging a friendship, will require a lot of contextual information about the person involved, the nature of the relationship, and the extent of the hurt feelings caused by the wrong, apologies do not work like that. While of course there is some room for subjective variation, it is possible to say in advance that an apology that does not contain the relevant elements (when it reasonably could have) will misfire, failing to bring about its characteristic normative effects. The appropriateness of the elements of apology to cases of wrongdoing is thus to an important degree independent of context. I will refer to this feature by saying that an unreserved apology is made up of a set of elements that are *canonical*.

These considerations count against a theory of apology that I will call the *reassurance view*. According to the reassurance view, apology gives us new practical reasons by providing evidence of psychological change in the wrongdoer. If it is often said that wrongdoing ruptures relationships, one version of the reassurance view sees this rupture as damaging the victim's confidence in their *own* moral standing. Another version, not necessarily incompatible with the first, sees it as damaging confidence or trust in the *wrongdoer*. The central feature of the reassurance view is thus that it sees wrongdoing as causing some harm, and apology as repairing that harm. Apology can repair the harm caused by wrongdoing, on the reassurance view, by the wrongdoer demonstrating a renewed commitment to moral standards. The distinctive moral effect of apology therefore consists in giving credible and practically relevant *evidence* of psychological change in the wrongdoer.⁶

However, as we can see from the preceding discussion, we can readily imagine a situation, like Missing Apology, in which the wrongdoer's actions have been such as to allow the victim to regain their confidence in their own moral standing and to allow the victim and others to place their trust in the wrongdoer again, but where no apology has been forthcoming. In such a situation, I

6 For an influential source of the reassurance view, see Murphy and Hampton, *Forgiveness and Mercy*. See also Gill, "The Moral Functions of an Apology." The reassurance view of apology might also fit well with Margaret Urban Walker's account of "moral repair" as "restoring or creating trust and hope in a shared sense of value and responsibility" (*Moral Repair*, 28). Similarly, Adrienne Martin puts it thus: "An apology typically includes saying one knows one flouted a legitimate norm and regrets it; the recipient needs to know the wrongdoer understands that he acted from an inadequate interpersonal commitment, if the recipient is to have reason to cease resenting that inadequate commitment" ("Owning Up and Lowering Down," 28). Note, however, that on Martin's account apology has performative as well as reassurance elements.

have suggested, it makes sense to think that there is normative work for apology to do that has not yet been done. However, the reassurance view does not have the resources to account for this. The job that, according to the reassurance view, it is the place of apology to carry out has in *Missing Apology* already been done.

Furthermore, I have argued that the elements required for a satisfying apology are canonical and largely determined independently of context, whereas if the reassurance view were correct, the acts that will provide the kind of evidence of psychological change a particular audience will need in order to be appropriately reassured would depend very much on the particular wrongdoer, the nature of the relationship, the particular victim, the expectations and biases of the audience, etc. Indeed, since evidence can be better or worse, the reassurance view gives us an account of the normative role of apology that is context dependent and scalar, thus failing to capture what is distinctive about the all-at-once nature of the closure hoped for in *Missing Apology*. A proponent of the reassurance view might argue, in rule-consequentialist fashion, that we have settled on a canonical set of elements as a kind of shorthand for signaling genuine remorse. But then we might query why performance of the shorthand would serve as good evidence of *genuine* remorse, rather than just a willingness to signal such remorse. At the very least, we would need a story about how the canonical elements could come to be the vehicle for genuine expressions of emotion.⁷

In trying to show that apology has an important normative function, the reassurance view takes it that this function is to repair the *harm* done by wrongdoing. More plausible, however, if we want to explain the distinctive role of apology in moral repair, is that the moral function of apology lies in addressing the fact that the victim has been treated *wrongfully*.⁸ The attractions of the reassurance view are clear, since repairing the harms of wrongdoing is indeed an important job and we can see in clear, nonmetaphorical terms how trust and confidence could be rebuilt after wrongdoing (albeit that it might be hard to do). By contrast, the idea of addressing past wrongs might seem less immediately urgent a task and perhaps even hopelessly metaphorical. However, if we want to explain why apology has a distinctive role among responses to wrongdoing, I will argue that we need to explore the idea that its function is to act on the normative situation *directly* and not simply as a source of evidence of the wrongdoer's state of mind.

7 For one version of such a story, see Pettigrove and Collins, "Apologizing for Who I Am," 144–48.

8 Cf. Hampton, "Correcting Harms versus Righting Wrongs."

II

In search of such an alternative, then, we might turn to the *performing deference view* of apology. The performing deference view understands the situation of wrongdoing as one in which the wrongdoer has subjugated the victim, degrading or demeaning them by treating them as lacking in moral status and thus as one whom it was permissible to treat in that way. This act creates harmful psychological effects, and, more saliently for the purposes of apology, it is a wrong done by the perpetrator against the victim. More metaphorically, perhaps, it is a disturbance in the normative order: an act that contravenes basic requirements of respect and consideration. On the performing deference view, the way to address the wrong and undo that normative disturbance is for the wrongdoer effectively to reverse that situation and to act out their subordination to the victim by means of apology. As Jeffrie Murphy puts it:

Wrongdoers attempt (sometimes successfully) to degrade or insult us; to bring us low; to say, “I am on high while you are down there below.” As a result we in a real sense *lose face* when done a moral injury—one reason why easy forgiveness tends to compromise self-esteem. But our moral relations provide for a ritual whereby the wrongdoer can symbolically bring himself low (or raise us up—I am not sure which metaphor best captures the point)—in other words, the humbling ritual of *apology*, the language of which is often that of *begging* for forgiveness. The posture of begging is not very exalted, of course, and thus some symbolic equality—necessary if forgiveness is to proceed consistently with self-respect—is now present.⁹

According to Murphy’s suggestion, then, the performance of the humbling ritual addresses not just the harm but the wrongdoing itself, and it does so by restoring the equality of relations that should have obtained and that the wrongdoing violated.

One potential advantage of Murphy’s view is that it explains why the body language of apology should exude deference and humility: the performance of subordination (somehow) restores moral equality. As far as Murphy tells us, however, the performance of subordination could be entirely insincere yet still do its normative work. Murphy says that “in the best of cases [apology] is likely to be a way of manifesting repentance.”¹⁰ However, he does not explain

9 Murphy, “Forgiveness and Resentment,” 28. Note that Murphy’s rich account can also be seen as a source of the reassurance view.

10 Murphy, “Forgiveness and Resentment,” 28.

how repentance can be manifested through an action with canonical features. Perhaps a proponent of the performing deference view could claim that the subordination of the wrongdoer is genuine only if the wrongdoer has internalized their inferior position and thus feels deferential as well as acting deferentially. Even with this addition, however, the view does not really explain how it is possible for such internalization to take place or how it can be that the wrongdoer authentically expresses the appropriate emotions through those symbolic actions.

Moreover, whether that strategy is plausible depends on the central question of whether the performing deference view is a good explanation of the characteristic normative effects of apology. While it is a strength of Murphy's view that it sees the characteristic role of apology as directly addressing the normative situation of wrongdoing, it is unclear how the fact that the wrongdoer performs a symbolic action would correct that situation. The wrongdoer, according to Murphy, goes through a ritual with a certain form, and let us grant that the form of the ritual is symbolically adequate to the nature of the wrongdoing (though I will dispute this below). Nevertheless, Murphy tells us nothing about why we should believe that such a performance could rationally be taken to bring about normative change.

In order to address this latter criticism, I suggest, we need to see apology not as a symbolic performance to no purpose but rather as a canonical action that brings about distinctive changes in the normative situation.¹¹ On this interpretation, the deferential behavior, along with the other canonical elements of apology, would be the vehicle through which the power to change the normative situation is exercised. The idea here would be of a certain sort of normative transfer: the situation of wrongdoing has brought about an imbalance in the proper distribution of respect, and the normative function of the symbolic performance involved in apology would be to rebalance things by taking away an excess and using it to restore a deficit.¹²

However, even with the performing deference view strengthened in this way, the key to the view remains its diagnosis of the initial moral situation addressed by apology: that wrongdoing brings low a moral subject who, because of basic

11 There are various ways to develop this idea, and in this paper, I do not commit myself to any particular framework. For instance, we might develop it as the idea of an Austinian performative. See Austin, "Performative Utterances." Or we could think of it as a normative power. See Raz, *Practical Reasons and Norms*, ch. 3; Owens, *Shaping the Normative Landscape*; and Bennett, "The Alteration Thesis." The idea that apology is a performative is most explicit in Helmreich's account of apology as "stance-taking" ("The Apologetic Stance"), though see also Martin, "Owning Up and Lowering Down."

12 See Bovens, "Apologies," who also uses this language.

equality, should never have been brought low, and who can be raised up again by the performance of subordination. The performing deference view falls apart if its diagnosis of the initial moral situation is faulty. For in that case, it would have no good explanation of why the performance of symbolically deferential behavior is necessary in the first place and thus no explanation of why that particular normative transfer is called for. As Jean Hampton brings out, however, Murphy's view about the moral situation addressed by apology is highly problematic. It seems to rest on the idea that the victim has in some way *actually been made less than equal* by virtue of the wrongdoing.¹³ Hampton argues that we should reject the idea that people can have their moral status altered by wrongdoing. She argues that we should rather hold the Kantian belief that moral status is unconditional, and that it is an implication of this Kantian belief that the victim cannot really have been lowered in their status. Murphy's view, according to Hampton, would have to rest on something more like a Hobbesian view on which moral status is a limited resource we compete and fight for, and where our ranking in the struggle can go up and down. Since this Hobbesian view is unacceptable, it follows that Murphy's view must be the wrong diagnosis of the normative situation that apology addresses, the wrong account of the normative effects of apology, and the wrong account of the symbolism of apology.

I have now looked at two influential recent accounts of the normative effects of apology. I argued that the reassurance view cannot explain why apology has canonical features, and that in seeking to explain apology as repairing harm rather than addressing wrongs, it cannot capture what is distinctive in apology. While the performing deference view represents an advance because it sees apology as addressing wrongdoing rather than harm, it misidentifies the need for apology and hence misinterprets its symbolism and distinctive normative effects. Furthermore, it provides no account of how symbolic performance can be both expressive and normatively powerful. In attempting to improve on the performing deference view, we need a better understanding of what it is for some action to be expressive of emotion, how expressions of emotion can involve canonical features, how the canonical features of expressions of emotion involve symbolism, and how the canonical features of an expression of emotion can become vehicles for the exercise of powers to alter the normative situation. Once we have a better understanding of these issues, we can then see how this might apply to the case of apology. I will start with an account of *expressive action*.

13 Hampton, "Forgiveness, Resentment and Hatred."

III

According to Jenefer Robinson, the core idea of “expression of emotion” is “a piece of behavior that manifests or reveals that emotion in such a way that we can not only infer from the behavior to the emotion but also perceive the emotion in the behavior.”¹⁴ “Ex-press” is here simply for something within to be pushed out. While I do not want to quibble about core meanings, the idea that I am interested in is somewhat different and closer to what we mean when we say that a piece of art is *expressively powerful*.¹⁵ For Robinson’s core idea of expression does not yet draw a distinction between *symptoms* of emotion and attempts to *give those emotions expression*. It does not distinguish, in other words, between behavior that merely *betrays* our emotions and behavior that is *expressive of* the emotion. There are many acts (as well as nonactions such as blushing and sweating) that may be caused by our emotions but are not expressive of them. Take a case in which I see a dark shape looming toward me as I walk through a darkened alleyway and my fear motivates me to put my hands up in a defensive position. Anyone witnessing this situation will be able to read my fear from my actions. Nevertheless, my behavior is not expressive of my fear: I do not *give* expression to my fear. By contrast, artworks can be attempts to give expression to emotions by the creation of objects or performances with properties that are expressively powerful in relation to some understanding of a situation. Furthermore, what is true of artworks can be extended to actions: when Christians kneel in church, what they do *is* (or can be) expressive of their sense that they are in the presence of a being whose worth is incomparably higher than their own.¹⁶

The point is not simply that the expression of reverence among churchgoers is conventional whereas the expression of fear in the alleyway is not. Rather

14 Robinson, *Deeper Than Reason*, 258.

15 For the account of expressive action developed here, see Bennet, “Expressive Actions,” “The Problem of Expressive Action,” and “How and Why to Express the Emotions.” The sense of “expression” delineated in the text is not identical with the type of expression explained in works in the philosophy of language such as Green, *Self-Expression*; and Davis, *Meaning, Expression and Thought*. More germane to my concerns is a debate in philosophy of action about expressive action, or action out of emotion, initiated by Hursthouse, “A-rational Actions.” While some action out of emotion is not expressive in the sense of being expressively powerful, I argue that there is a wide class of expressively powerful actions and, furthermore, that their expressive power is a good (rational) explanation of why we do them. See also, e.g., Betzler, “Expressive Actions”; and Döring, “Explaining Action by Emotion.”

16 As it happens, I am not a Christian, but I hope readers Christian and non-Christian can understand the point of the example.

the key point is that whereas the defensive posture is simply *caused* by my fear, the expression of reverence is expressively *powerful* in relation to the content of the feelings experienced by the churchgoers (in the way that artworks can be expressively powerful in relation to such content). In other words, the form taken by the expression of reverence (that of kneeling or lowering oneself) can be seen as *reflecting* or *capturing* the content of the relevant attitudes—that is, the perception of the incomparably higher worth or value of the Divine. By contrast, when I assume a defensive posture out of fear, I am simply trying to defend myself. There *could* be acts that are expressive of fear: for instance, if before I give a talk at an important conference, I act as though my legs have gone to jelly. But the point is that not all of those acts that betray, or are caused by emotion, are expressive of that emotion in the sense in which we are interested here. What is distinctive of acts that are expressive of emotion is that they are expressively powerful by virtue of the fact that the form they take reflects important elements of the situation at which they are directed.

If this is correct, then we have established that behavior being caused by emotion is not *sufficient* for that behavior to be expressive of that emotion—not in the sense of being expressively powerful in relation to the content of that emotion. However, neither is being caused by the emotion *necessary* for the act to be expressively powerful. Acting as though my legs have gone to jelly can be a powerful expression of fear even if I am not actually feeling fear at the time I engage in this action. Even if I am feeling perfectly confident about my talk, for instance, I might do the wobbly legs routine in order to indicate that I understand that this is an important event and the kind of thing it is quite appropriate to be nervous or fearful about. Perhaps I do this in order to show solidarity with other speakers who are feeling more nervous than I am. Thus, we have the possibility that an expressive act comes to take on social meaning that does not depend on the motivations of the agent.¹⁷

I now want to argue that an act is expressive of some emotion insofar as it is a powerful *symbol* of that emotion.¹⁸ What I mean by “symbol” here is not that the behavior simply *denotes* fear according to some conventional scheme of reference (as an ox in a painting of a saint may denote Luke, and a winged lion, Mark) but rather, to adopt Nelson Goodman’s distinction between types of reference, that it (metaphorically) *exemplifies* it.¹⁹ To exemplify, for Goodman, is to refer to a property by possessing it, as a sample of cloth refers to the

17 Anderson and Pildes, “Expressive Theories of Law.”

18 Bennett, “Expressive Actions.”

19 Goodman, *Languages of Art*, 85. See also the discussion in Eldridge, *An Introduction to the Philosophy of Art*, ch. 4.

cloth itself. A property is expressive of something, on Goodman's view, when it metaphorically exemplifies it. Thus, I will say that some action (or other vehicle) is expressive of an emotional state when it symbolizes those features of the situation that that particular emotion makes salient. That is, it symbolizes the awe- or fear- or remorse- or joy-worthy features of the situation: those features that call for that emotion. And it symbolizes those features when it is such that we can (metaphorically) see those salient features *in* (or exemplified by) that action. For instance, in the kneeling we can see the Christian's perception of their situation as one in which their significance is dwarfed by the incomparable worth of the Almighty.

What makes something expressive of an emotion is therefore not a causal link that it bears to an emotional state but rather its *expressive properties*. And if we follow Goodman, we will say that the properties of an action are expressive when they bear a relation to the referent that is not conventional but rather a matter of intelligible gestalt—of what we can intelligibly see or construe as a telling metaphor for the referent. On this sense of “expressive,” it is possible to engage in expressive actions without experiencing the emotion at the time. An action can be expressive of an attitude without being used in that instance by an agent to express their own attitude.²⁰ Nevertheless, there is something about the form of the action—for instance, its symbolic properties—that makes it particularly appropriate as an expressive vehicle for that emotion. Furthermore, as we will see below, for an action to be expressive in this sense, it must be the case that although the actual underlying presence of emotion is neither necessary nor sufficient for the act to be expressive, it is nevertheless quite intelligible—indeed normal—for people in the grip of an emotion to give expression to it by engaging in those actions.

IV

Why would one engage in expressive actions? Action that is expressive is not aimed, in the first instance, at altering the material situation or bringing about some further end. But this does not necessarily mean that it is pointless. Rather, if there is some point to engaging in expressive action, it is what we might call a backward-looking one: it lies in *marking* the situation as important in some distinctive way and attempting to *do justice to* it.²¹ The point of expressive actions is simply to *acknowledge* or *recognize* their significance. This phenomenon of

20 Thus, we also have the possibility of nonsincere expressions of emotion—that is, of acts that present themselves either as (falsely) indicating the presence of the emotion or as indicating the agent's sense of the appropriateness of the emotion to that situation.

21 Bennett, “Expressive Actions.”

“doing justice” to one’s situation (through expressive actions) is important in our life because it enables us to isolate that situation from the ongoing rush of “one damn thing after another” and allows our attitudes to that situation to themselves become an object of scrutiny. Having an action that resonates with those attitudes and in which we can see those attitudes represented allows us to dwell on the situation and what is salient in it. Perhaps we could even say that such action is a vehicle through which we can dwell in and with what is salient in the situation for an intense period of reflection. Expressive action therefore allows us to mark certain situations or events as pivotal or as otherwise out of the ordinary—as something to which special attention must be paid—through actions that resonate with our sense of why those situations are salient.

Expressive actions, once one starts to look for them, are common. They are the kind of thing we do when we welcome or take leave, when we mourn or celebrate or commiserate, when we thank, or (perhaps) when we blame. Sometimes the same action, broadly described, can be expressive of quite different emotions, and it is the context and the emotional tone of the performance—or the way in which the action is performed—that make it clear which emotion is being expressed. For instance, embracing a loved one symbolizes and is expressive both of being pleased to see them and of being sorry to see them go. In both cases, the act of holding them close is expressively powerful in relation to the past or future in which they are far away. Thus, embracing can be carried out in various ways to capture regret or delight. Sometimes we get it wrong and feel that the departing embrace was too cursory, or that we brushed the departure off lightly, that the embrace did not at all capture what we were feeling (or feel now that we are away from the immediate emotional pressure of the situation).

Expressing one’s emotions can therefore be a complex and creative affair. On the one hand, this is because of the complexity of the situations we find ourselves in, and the multiple aspects that may call for different emotions, such that the dominant saliences are unclear, at least prior to deliberation. On the other hand, it is because each of us may bring an individual style to our expressiveness. Expressive acts have to fit in with the ways of acting that are characteristic of the agent. It can take a while for a person to find their expressive style, and some people are better at it than others. However, even when the emphasis is on finding our own way of expressing what we feel about a situation, we do so against shared background understanding of paradigm scenarios that (perhaps relative to some community of interpreters) provide canonical understandings of which actions are expressively powerful in relation to which situations.²²

22 De Sousa, “The Rationality of Emotions.”

When Christians enter a church, or a parent says goodbye to the child who is going away, they might look, I have suggested, for some action to do justice to the way they feel. But they will not have to search *too* far. They are not beginning the search *ex nihilo*. Rather, in coherentist fashion, they take up some action that seems right to them on the basis of understandings already there in the culture. Again, this is not simply a question of adopting the local conventions. It is a matter of finding metaphors or symbols compelling and powerful on the basis of the way in which they fit with other aspects of one's inherited background of beliefs and values and vocabulary of other expressive actions. Once we see that expressive actions are grounded in an understanding of the power of certain symbols and metaphors, and that these understandings can be shared and embedded in a culture, there is no difficulty in seeing how it could be that individuals can express their own emotions through canonical actions.

We are now in a position to see how it might be at least possible that an apology could be expressive even though it deploys the canonical actions through which it can generate regular normative effects. The fact that the expressive significance of some acts has become common currency in a culture does not mean that we cannot perform them as authentic expressions of emotion. We can imagine the practice of apology altering and developing if agents came to experience the instituted understandings as inappropriate and unsatisfying, or merely ritualistic. It has to be the case that people can use these socially instituted (or collectively developed) forms for the genuine expression of emotion. A practice might die or be radically altered if an adequate expressive vehicle cannot be found. I take it, however, that, for many of us at least, this is not the case with our practice of apology; it is an interesting feature of our—in many respects highly diverse—society, and the patterns of socialization at work in it, that the very same actions necessary to do the normative work of apology can also be experienced as vehicles for the authentic expression of remorse.

V

It is an interesting question how social and historical development might bring it about that the features of expressive actions could become canonical and under what conditions there might be social pressure toward unanimity. Perhaps people converge on a given set of symbols because of their inherent aptness (given a background of other widely accepted beliefs, values, and symbols); or perhaps the convergence is to be explained by the influence of an established religion that preaches a sacrament of penance; or perhaps there is

something of truth in both of these explanations. However, we might imagine that the conception of expressive power on which a community converges feeds into its conception of ritual observance, delivering a shared sense of the actions that are necessary to mark and do justice to the major normative transitions of human existence. Given this convergence, it may be plausible that, over generations, the form of the ritual widely judged to be appropriate starts to mold our emotions and perceptions, until it is the ritual that comes to form the narrative arc of the emotion, the lack to which it responds and the satisfaction it seeks.

Furthermore, we can perhaps now speculate that it would be plausible that if the expressive action of apology did become canonical, then it could also become intertwined with a community's normative understandings and come to take on the significance of a power through which regular normative effects are brought about. If there are such powers—or if a particular culture develops the idea of such powers—it might come to seem important that the form taken by those actions be not simply arbitrary. After all, the action is an important one that brings about weighty normative changes, and it might seem that the action should reflect that significance. Thus, it might seem that the power could only be exercised by acts that are particularly fitting to the normative situation. Perhaps we could interpret this as the idea that the form taken by those actions should be in some way *continuous* with the normative effects being brought about. (An archaic example of such a symbolically adequate power might be where a courtier bows or kneels to a monarch before approaching them for an audience, assuming a deferential posture in order thereby to make such an approach permissible.)

Furthermore, it might come to seem necessary that one cannot exercise such a power *lightly*, but rather that one has to be in a state of awareness of the significance of the power being exercised. If this were the case, then it would be not only the *form* of one's actions but also the *spirit* in which one does them that has to fit with the gravity of the situation. In such a case, we might say, one changes one's normative situation by means of an action that expressively recognizes the significance of the normative distance to be traversed. If this seems plausible, it suggests that for some powers, we should expect the action that exercises the power not to be one that is arbitrarily specified by convention, but rather one that is expressively adequate. The action has to be expressively adequate both in the sense that its form has to correspond to the normative situation being altered and in the sense that the person who exercises the power has to do so with an awareness of the gravity of the situation. In order to exercise the power, one has to act in the way that someone would who appropriately recognizes the gravity of the situation they are in.

VI

How does this account of expressive actions apply to apology? A full answer to this question would have to explain what kind of expressive action apology is and what characteristic normative effects are brought about by its felicitous performance. It would need to explain how the form taken by the actions required by apology is fitting to the situation of wrongdoing. And it would have to explain how those fitting actions are also fitting to the normative effects apology brings about. Defending such an answer in full would require a further paper. However, it is possible to provide a brief sketch of how an account of apology as an expressive action might be developed.

To start with, consider that a common “protesting” response to wrongdoing consists in a refusal to engage in normal relations with the wrongdoer (perhaps until such time as they put things right). It might be asked why we react to wrongdoing this way, with what P. F. Strawson calls a “partial and temporary withdrawal of goodwill.”²³ Is it simply instinct, or morally arbitrary socially constructed behavior? Or is it in some way fitting to the situation? We can explain the intuition that such behavior is fitting by understanding such distancing as an expressive action in the terms just outlined. The distancing is expressive of emotions of condemnation when it is performed as a compelling way of doing justice to the salient features of the situation of wrongdoing. The distancing symbolizes (metaphorically exemplifies) the normative situation of an agent who has violated a fundamental norm of the moral community to which they belong as a self-governing member. On the basis of such membership, the agent would normally be due certain distinctive marks of respect and recognition. Distancing is expressively powerful because it consists in a partial and temporary withdrawal of that respect and recognition, which is carried out because it marks the fact that the wrongdoer has done something that members of the moral community should be committed not to doing. The wrongdoer is not expelled from the community. Expulsion would be entirely the wrong symbolism. They are still within the community, fully subject to its norms, and deserving of the recognition due to its members. But they have acted in a way that a member of the community should have seen as impermissible. Hence, withdrawing recognition is a way of *reaffirming* the wrongdoer’s membership in the community despite their wrongdoing and, as such, is an apt symbol to capture the moral situation in which the wrongdoing has placed the wrongdoer.²⁴

23 Strawson, “Freedom and Resentment,” 77.

24 For the account sketched here, see Bennett, “The Varieties of Retributive Experience” and *The Apology Ritual*, ch. 5.

What we have talked about so far is someone distancing themselves from *another person* as a way of protesting what they have done. However, if apology is an expressive action, then it must be expressive of an attitude toward *oneself* and one's *own* situation. But we can straightforwardly extend the account just given to explain the canonical features of apology. In the case of remorse or self-blame, the expressively powerful forms of behavior have to do with distancing *from oneself*. One withdraws normal relations, as it were, from oneself. It is this self-withdrawal that accounts for the deferential and humble posture of the remorseful, and the penitential willingness to renounce benefits that would otherwise have been one's due and undertake tasks that would not otherwise have been one's duty. Unlike the performing deference view, which explains the appropriateness of deferential behavior by reference to the need to restore the victim's moral equality, the view I am arguing for is not committed to the claim that the victim has initially been brought low. What is rather at issue is the offender's having committed a wrong that distances them from the moral community.

In engaging in deferential, self-denying behavior, the offender *joins with others* in distancing from the wrongdoer. Thus, as long as it is the wrongdoer's own, authentic remorse being expressed through these symbols of self-withdrawal, apology is expressive not just of distance from their own wrongdoing but also of the fact that the wrongdoer is no longer at odds with the moral community. The wrongdoer has taken up the attitude of condemnation shared by other members of the moral community, seeing their action as to be repudiated, to be dissociated from, and has striven to do justice to that attitude by altering their treatment of themselves accordingly, through penitential actions and postures of humility.

Nevertheless, the fact that apology expresses return to the community is not the end of the matter: the wrongdoer is not back on equal terms, but rather on probation. What apology does is therefore a beginning of the process of return, rather than its end. It is expressive of a commitment to stay within the community. The offender whose postapologetic behavior is not expressive of this probationary status risks undermining the meaning of their apology. Thus, it is typically expressively inappropriate simply to return to a carefree demeanor too quickly after apologizing for a significant wrong.

We can now extend our sketch to show how it might be argued that there is a relation of fit between this symbolism and the normative effects that apology purports to bring about. Apology, it might be suggested, has two characteristic normative effects. First, it alters the normative situation so that others can resume normal relations with the offender without moral compromise. And second, it gives the addressee of the apology special rights that the offender

answer to them for their conduct in relation to such desistance: a kind of special oversight authority over that conduct. These normative effects echo the expressive properties of the action whereby the offender dissociates from the earlier self who performed the wrongful action. The vehicle of apology—the actions necessary for the exercise of this power—are not conventional and arbitrary but rather fitting to the situation in which they are exercised. It is the absence of an act with these normative effects that explains the sense of dissatisfaction articulated in *Missing Apology*.²⁵

The account I have sketched here would explain how apology can be expressive of remorse, how it can be restorative, and why it involves some commitment regarding future behavior. If we want to give a name to this account of apology as an expressive action, I would describe it as a theory of apology as *dissociation*.²⁶ In apology, one performs actions that are expressive of dissociation from one's past wrong and reassociation with the moral community. And when those expressive actions express one's genuine remorse, they could, if the view sketched here is plausible, have the normative effect of dissociating oneself from one's past wrong and returning one to the community, albeit with a probationary status, thereby allowing others to resume normal relations without moral compromise.

VII

In this paper, I have looked at how we might rationalize our practice of treating apologies as bringing about distinctive normative effects. Since apology involves a set of canonical actions but is also expressive of remorse, I argued that it is necessary to understand it as an expressive action. Having explained what expressive actions are, I sketched a view on which the expressively powerful, or expressively adequate, way to do justice to wrongdoing is *dissociation*—that is, not to treat the offender normally, and specifically to withdraw

25 I deny that apology can have its backward-looking normative effect of making uncompromised resumption of relations possible unless it is an expression of remorse. However, perhaps an apology can have its forward-looking effect of transferring oversight authority *without* being given sincerely. That is, it may be plausible to say that even the person who gives an apology grudgingly has, simply by making the apology, thereby created a new normative relation to the addressee in virtue of which they can appropriately be called to account by the addressee for failing to take relevant steps toward desistance.

26 For some further thoughts about dissociation, see Bennett, "Complicity and Normative Control." The notion of dissociation through apology is also appealed to in Hieronymi, "Articulating an Uncompromising Forgiveness."

the respect to which they would otherwise have been entitled. Dissociation, I suggested, is what the offender does to himself when he apologizes.²⁷

University of Sheffield
c.bennett@sheffield.ac.uk

REFERENCES

- Anderson, Elizabeth, and Richard Pildes. "Expressive Theories of Law: A General Restatement." *University of Pennsylvania Law Review* 148, no. 5 (May 2000): 1503–75.
- Austin, J. L. "Performative Utterances." In *Philosophical Papers*, edited by J. O. Urmson and G. Warnock, 233–52. Oxford: Clarendon Press, 1961.
- Barnum-Roberts, Brooke Natalie. "Apologizing without Regret." *Ratio* 24, no. 1 (March 2011): 17–27.
- Bennett, Christopher. "The Alteration Thesis: Forgiveness as a Normative Power." *Philosophy and Public Affairs* 46, no. 2 (Spring 2018): 207–33.
- . *The Apology Ritual: A Philosophical Theory of Punishment*. Cambridge: Cambridge University Press, 2008.
- . "Complicity and Normative Control." *Monist* 104, no. 2 (April 2021): 182–94.
- . "Expressive Actions." In *The Expression of Emotion: Philosophical and Legal Perspectives*, edited by Catherine Abell and Joel Smith, 73–94. Cambridge: Cambridge University Press, 2016.
- . "How and Why to Express the Emotions: A Taxonomy of Emotional Expression with Historical Illustrations." *Metaphilosophy* 52, no. 5 (October 2021): 513–29.
- . "The Problem of Expressive Action." *Philosophy* 96, no. 2 (April 2021): 277–300.
- . "The Varieties of Retributive Experience." *Philosophical Quarterly* 52, no. 207 (April 2002): 145–63.

27 Versions of this paper were presented to audiences at Stockholm, Munich, and Essen, and to the Society of Applied Philosophy Annual Conference in Cardiff. I would like to thank those who gave me useful feedback at those events, in particular: Monika Betzler, Susanne Boshammer, Andreas Brekke Carlsson, Peter Chau, Garrett Cullity, Helen Frowe, André Grahle, Charles Griswold, Oliver Hallich, Richard Healey, Bennett Helm, Sofia Jeppsson, Julian Jonker, Ellie Mason, Ben Mathieson, Leonhard Menges, Per Milam, Jennifer Page, Glenn Pettigrove, Linda Radzik, Stefan Riedener, Peter Schaber, Adam Slavy, and Bill Wringe. I have also benefited from very helpful input from referees.

- Betzler, Monika. "Expressive Actions." *Inquiry* 52, no. 3 (2009): 272–92.
- Bovens, Luc. "Apologies." *Proceedings of the Aristotelian Society* 108, no. 1 (October 2008): 219–39.
- Cohen, Andrew. "Vicarious Apologies as Moral Repair." *Ratio* 30, no. 3 (September 2017): 359–73.
- Davis, Paul. "On Apologies." *Journal of Applied Philosophy* 19, no. 2 (2002): 169–73.
- Davis, Wayne A. *Meaning, Expression and Thought*. Cambridge: Cambridge University Press, 2002.
- de Sousa, Ronald. "The Rationality of Emotions." In *Explaining Emotions*, edited by Amélie Oksenberg Rorty, 127–52. Berkeley: University of California Press, 1980.
- Döring, Sabine A. "Explaining Action by Emotion." *Philosophical Quarterly* 53, no. 211 (April 2003): 214–30.
- Eldridge, Richard. *An Introduction to the Philosophy of Art*. Cambridge: Cambridge University Press, 2003.
- Gill, Katherine. "The Moral Functions of an Apology." *Philosophical Forum* 31, no. 1 (Spring 2000): 11–27.
- Goodman, Nelson. *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis: Hackett, 1976.
- Govier, Trudy, and Wilhelm Verwoerd. "The Promise and Pitfalls of Apology." *Journal of Social Philosophy* 33, no. 1 (Spring 2002): 67–82.
- Green, Mitchell. *Self-Expression*. Oxford: Oxford University Press, 2007.
- Hampton, Jean. "Correcting Harms versus Righting Wrongs: The Goal of Retribution." *UCLA Law Review* 39 (1992): 1659–702.
- . "Forgiveness, Resentment and Hatred." In Murphy and Hampton, *Forgiveness and Mercy*, 35–87.
- Helmreich, Jeffrey S. "The Apologetic Stance." *Philosophy and Public Affairs* 43, no. 2 (Spring 2015): 75–108.
- Hieronimi, Pamela. "Articulating an Uncompromising Forgiveness." *Philosophy and Phenomenological Research* 62, no. 3 (May 2001): 529–55.
- Hursthouse, Rosalind. "Arational Actions." *Journal of Philosophy* 88, no. 2 (February 1991): 57–68.
- Joyce, Richard. "Apologizing." *Public Affairs Quarterly* 13, no. 2 (April 1999): 159–73.
- Lazare, Aaron. *On Apology*. Oxford: Oxford University Press, 2004.
- Martin, Adrienne M. "Owning Up and Lowering Down: The Power of Apology." *Journal of Philosophy* 107, no. 10 (October 2010): 534–53.
- Murphy, Jeffrie G. "Forgiveness and Resentment." In Murphy and Hampton, *Forgiveness and Mercy*, 14–34.

- Murphy, Jeffrie G., and Jean Hampton, *Forgiveness and Mercy*. Cambridge: Cambridge University Press, 1988.
- Owens, David. *Shaping the Normative Landscape*. Oxford: Oxford University Press, 2012.
- Pettigrove, Glen, and Jordan Collins. "Apologizing for Who I Am." *Journal of Applied Philosophy* 28, no. 2 (May 2011): 137–50.
- Raz, Joseph. *Practical Reasons and Norms*. 2nd ed. Oxford: Oxford University Press, 1999.
- Robinson, Jenefer. *Deeper Than Reason: Emotion and Its Role in Literature, Music, and Art*. New York: Oxford University Press, 2005.
- Smith, Nick. "The Categorical Apology." *Journal of Social Philosophy* 36, no. 4 (Winter 2005): 473–96.
- . *I Was Wrong: The Meanings of Apologies*. Cambridge: Cambridge University Press, 2008.
- Strawson, Peter. "Freedom and Resentment." In *Free Will*, edited by Gary Watson, 59–80. Oxford: Oxford University Press, 1982.
- Tavuchis, Nicholas. *Mea Culpa: A Sociology of Apology and Reconciliation*. Stanford, CA: Stanford University Press, 1991.
- Walker, Margaret Urban. *Moral Repair: Reconstructing Moral Relations after Wrongdoing*. Cambridge: Cambridge University Press, 2006.

AGAINST BEING FOR

James L. D. Brown

EXPRESSIVISM has always been a shifting target. Throughout its history, expressivists have denied such things as the truth-aptness of normative sentences, the existence of normative facts and properties, and the cognitive status of normative attitudes. But in each case, expressivists have come to embrace what was previously denied. To a large extent, this is a product of so-called creeping minimalism, where expressivists embrace deflationary interpretations of notions like *truth*, *fact*, and *property*.¹ But some expressivists aim to accommodate robust conceptions of these notions within their theory. For instance, it has been argued that expressivists can be correspondence theorists about truth, and that expressivists should identify normative properties with natural properties.²

This paper focuses on normative *belief*. Specifically, it examines whether expressivists can develop a robust theory of normative belief compatible with expressivism. Initially, expressivists denied that normative judgments are beliefs. But this view faces the awkward fact that normative belief-talk is ubiquitous in ordinary language. To accommodate this, many expressivists distinguish between robust and minimal senses of “belief,” embracing minimal normative beliefs while rejecting robust normative beliefs. However, it is not simply that we use the same language to *talk* about normative attitudes. Normative attitudes possess many of the distinctive features of belief *tout court*.³ So we are left with a view according to which there are two classes of mental states that *just so happen* to have all the same core properties despite having a completely different nature. And one might worry that this looks suspect.

Related worries have led some to question whether expressivists might instead maintain that normative claims express robust beliefs, but that such

1 See Dreier, “Meta-Ethics and the Problem of Creeping Minimalism.”

2 For the former claim, see Ridge, *Impassioned Belief*. For the latter claim, see Gibbard, “Normative Properties”; and Bex-Priestley, “Expressivists Should be Reductive Naturalists.”

3 See Horgan and Timmons, “Cognitivist Expressivism.”

beliefs are nondescriptive.⁴ This can be done by providing a nondeflationary, nondescriptive theory of belief *tout court* which can distinguish between nondescriptive normative beliefs and ordinary descriptive beliefs. For instance: Horgan and Timmons offer a commitment-based theory; Köhler offers a functionalist theory; I have offered an interpretationist theory; Gibbard can be read as providing a sententialist theory; one might construct such a theory within a cognitive act view of propositions; and Schroeder offers a theory of belief as being for.⁵

This paper contributes to the task of assessing this expressivist approach to normative belief by critically examining the last of these views: Schroeder's theory of belief as being for. After outlining the view (section I), I raise a challenge for it (section II), ultimately concluding that we should reject the view as it stands.

I

Schroeder's theory of belief as being for is offered as an instance of a more general nondescriptivist framework for explaining belief. It will therefore be helpful to begin by outlining the more general framework. Schroeder begins by differentiating two theoretical roles that propositions have traditionally been posited to play.⁶ The first role is that of being the objects of attitudes and the primary bearers of truth and falsity. The second role is that of carving up the world, where propositions correspond to distinctions in reality. He calls the entities that play the first role *propositions* and the entities that play the second role *representational contents*. If beliefs are essentially descriptive, then it is natural to assume that propositions just are representational contents. However, if beliefs are not essentially descriptive, then we can reject this assumption. Instead, Schroeder proposes that there are two distinct classes of entities corresponding to each role, and that propositions and representational contents are "two different sorts of thing."⁷

Schroeder then proposes that while all beliefs involve an agent being related to a proposition, descriptive beliefs also involve the agent being related to a

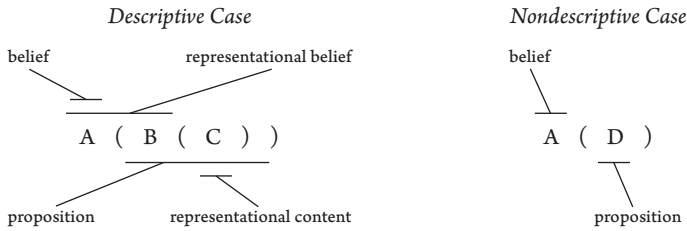
4 For an overview of these worries, see Schroeder, "Two Roles for Propositions."

5 Respectively, see Horgan and Timmons, "Cognitivist Expressivism"; Köhler, "Expressivism, Belief, and All That"; Brown, "Interpretative Expressivism"; Gibbard, *Meaning and Normativity*; Brown, "Expressivism and Cognitive Propositions"; Schroeder, *Being For* and "Two Roles for Propositions."

6 Schroeder, "Two Roles for Propositions," 418.

7 Schroeder, "Two Roles for Propositions," 421.

representational content. Diagrammatically, Schroeder suggests the following picture:



So in the nondescriptive case, a belief is simply a relation to a proposition. But in the descriptive case we can carve up belief in two ways. Specifically, belief can be carved up as the relation $A(_)$ to a proposition $B(C)$ or it can be carved up as a relation $A(B(_))$ to a representational content C .⁸

Lest the descriptive case seem implausibly complex, Schroeder offers the following comparison.⁹ Consider the state of *being about to go to Paris*. On one way of carving up this state, it consists in an agent standing in the relation of *being about to* and the act type of *going to Paris*. But on another way of carving up the state, it consists in an agent standing in the relation of *being about to go to* and the city Paris. So the single state of *being about to go to Paris* consists in two distinct relations to two distinct objects. This example exhibits the same structure of the descriptive case in the diagram above. Here, “ $A(_)$ ” denotes *being about to*, “ $B(_)$ ” denotes *going to*, “ $A(B(_))$ ” denotes *being about to go to*, and “ C ” denotes Paris. So there is nothing inherently problematic or unusual with a state having this structure.

Schroeder’s theory of belief as being for is then offered as an implementation of the more general framework provided by the diagram. Although Schroeder expresses more confidence in the general framework, my focus here will be the theory of belief as being for.¹⁰ But it is worth highlighting that expressivists can embrace normative beliefs and propositions without adopting this framework. A simpler framework would maintain that all beliefs involve a single relation between an agent and a proposition, where some but not all propositions are representational contents.¹¹ On this view, descriptive beliefs are relations to propositions that are representational contents, and nondescriptive beliefs are

8 Schroeder, “Two Roles for Propositions,” 422.

9 Schroeder, “Two Roles for Propositions,” 421

10 Schroeder, “Two Roles for Propositions,” 424.

11 See, for instance, Brown, “Interpretative Expressivism.”

relations to propositions that are not representational contents.¹² So it is not clear what advantages are gained by adopting Schroeder's general framework. However, this point will not matter for the discussion that follows, so I will not pursue it any further here.

Returning to Schroeder's view, he postulates the attitude of *being for* to play the A-role of belief within his framework. An agent has the attitude of being for when they are *for* something. As a first approximation, what one is for is a certain kind of property. So being for is a relation that holds between agents and properties. More specifically, when one is for something, then other things being equal, one *does* that thing.¹³ So being for relates agents to possible *act types*, broadly construed. We can then define the *being for* relation as the state "whose functional role is to lead one to acquire that property, other things being equal."¹⁴ If believing some proposition consists in being for some property, then we can say that propositions are the properties that one is for when one believes that proposition. If the relevant class of properties are act types, then identifying the proposition *p* will involve specifying what an agent is typically motivated to do when she believes *p*.¹⁵

This general account of belief is perfectly suited for an expressivist theory of normative belief. Expressivists think that normative thought is essentially directive, in the sense of being action guiding or attitude governing. If believing consists in being for, then it turns out that all beliefs are essentially directive. The directive nature of normative belief is then just a particular instance of the directive nature of belief in general. So if we suppose that normative expressions like "wrong" express noncognitive attitudes like *disapproval*, we can identify the belief that murder is wrong with *being for disapproving of murder*. If believing just is being for, then the proposition that murder is wrong just is the property of disapproving of murder. In terms of the diagram, *A* is being for, and *D* is the act type that one is for when one believes *D*.

With the nondescriptive case explained, we now need to explain the descriptive case. If believing is being for, and being for involves being motivated to do something, what kind of things are we motivated to do when we have a descriptive belief? To explain this, Schroeder introduces the notion of *proceeding as if*, which relates agents to representational contents. In terms of the diagram,

12 As well as being simpler, this framework seems to preserve all the explanatory advantages Schroeder ascribes to his framework; see Schroeder, "Two Roles for Propositions," 422–23.

13 Schroeder, *Being For*, 84.

14 Schroeder, "Two Roles for Propositions," 424.

15 To explain negated descriptive beliefs, the theory actually requires propositions to be *pairs of entailing properties*; see Schroeder, *Being For*, 95–100. However, this detail will not be important in what follows, so I will stick to the simplified version for ease of exposition.

this is relation *B*. Descriptive belief therefore consists in *being for proceeding as if p*, where “*p*” denotes a representational content. Thus, for instance, believing that the cat is on the mat consists in *being for proceeding as if the cat is on the mat*, where “*the cat is on the mat*” denotes a representational content.

What is it to proceed as if *p* and to be for proceeding as if *p*? Schroeder provides the following answer:

On my best gloss, to proceed as if *p* is to take *p* as settled in deciding what to do. So being for proceeding as if *p* is being for taking *p* as settled in deciding what to do. Assuming that being for has the motivational property that someone who is for *a* will tend to do *a*, other things being equal, it follows that someone who believes that *p* will tend to proceed as if *p*, other things being equal. That is, it follows that she will tend to treat *p* as settled in deciding what to do.¹⁶

At a descriptive level, Schroeder’s suggestion seems plausible enough. When I believe that the cat is on the mat, other things being equal, I do proceed as if the cat is on the mat, in that I take this fact as settled in deciding to avoid walking on the mat, to look for the cat on the mat, and so on. So if believing is being for, then the proposition that the cat is on the mat is the property of *proceeding as if the cat is on the mat*.

Schroeder’s theory of belief as being for thus simultaneously provides a novel account of belief and belief contents suitable for expressivists. The account is fully general, applying to all beliefs, but it is designed to distinguish between descriptive beliefs that involve some kind of ontological commitment and normative beliefs that do not. Clearly, then, a central criterion of success will be whether the theory successfully does this. It is this aspect of the theory that I will now challenge in the next section.

II

The notion of *proceeding as if* plays a central role for Schroeder in distinguishing descriptive from nondescriptive belief. However, it is unclear whether it can play this role because the notion applies just as much in the nondescriptive case as in the descriptive. If I believe that the cat is on the mat, then a plausible description of my state is that I am for proceeding as if the cat is on the mat. However, it is no less plausible to describe my belief that murder is wrong in terms of my being for proceeding as if murder is wrong. After all, other things being equal, if I have this belief, then I will take the wrongness of murder as

¹⁶ Schroeder, *Being For*, 93–94; see also Schroeder, “Two Roles for Propositions,” 426.

settled in deciding what to do. So, intuitively, the notion of proceeding as if fails to distinguish between descriptive and nondescriptive beliefs.

One response to this objection would be to point out that “proceeding as if” is defined as a relation between (1) states of being for and (2) representational contents. If we reject the existence of normative representational contents, then there is nothing that can fill the place of 2, and so there are no normative instances of proceeding as if. That is, because “murder is wrong” expresses only a proposition and no representational content, and because by definition proceeding as if is a relation between states of being for and representational contents, and not between states of being for and propositions, there is no possible state of being for proceeding as if murder is wrong.

However, although it is possible to define “proceeding as if” in this way, the stipulation that it applies only to representational contents is not grounded in the functional characterization of the attitude. If we take “proceeding as if” as “a shorthand for the general relation of taking something as settled in one’s deliberative activity,” as Schroeder suggests, then the notion more plausibly applies to propositions, even nondescriptive propositions, because it is clear that I can take the proposition that murder is wrong as settled in my deliberative activity.¹⁷

Perhaps one could amend the functional characterization to range over representational contents explicitly, so that “proceeding as if” is shorthand for the relation of taking *representational contents* as settled in one’s deliberative activity. However, it is unclear that we have any grasp of this functional role over and above that of the original functional role. Moreover, explicitly ranging over representational contents in this way looks more like a criterion of adequacy that the functional role of proceeding as if must meet rather than an explanation of what this functional role is actually like, such that it only ranges over representational contents. So we have not yet been given any psychological distinction to go along with our distinction between $\langle \text{pai } p_{\text{proposition}} \rangle$ and $\langle \text{pai } p_{\text{representational content}} \rangle$.

Indeed, insofar as proceeding as if has its home in deliberative activity, then even in the descriptive case it seems more plausible that proceeding as if is a relation to propositions and not representational contents. For deliberative activity constitutively involves propositional acts and attitudes, and so the most natural way of characterizing this activity will appeal to the propositional contents of those acts and attitudes, including in the descriptive case. Moreover, it would be strange if the same activity took a different object in the descriptive case even though the same kind of object (propositions) is also available.

17 Schroeder, “Two Roles for Propositions,” 426.

However, this might suggest another response to the objection. While Schroeder explains descriptive belief in terms of proceeding as if, he is aware that one can also describe normative beliefs in terms of proceeding as if.¹⁸ Schroeder suggests that we explain the difference in terms of what proceeding as if *consists in* in each case. Thus, an expressivist might claim that while it is correct to describe the belief that murder is wrong as being for proceeding as if murder is wrong, proceeding as if murder is wrong *just is* disapproving of murder. By comparison, while it is also correct to describe the belief that the cat is on the mat as being for proceeding as if the cat is on the mat, proceeding as if the cat is on the mat *just is* proceeding as if a certain representational content is true.

However, this response faces the same objection as before. For we do not have any account of what it is to proceed as if some representational content is true. Again, insofar as we understand what it is to proceed as if p , it more plausibly applies to propositions and not representational contents. Further, if the very same relation applied to propositions and representational contents, then this would over-generate beliefs. For every descriptive state picked out by the expression “being for proceeding as if p ,” “ p ” would ambiguously denote a proposition or a representational content. Given Schroeder’s assumption that these are two distinct kinds of entity, it follows that there are two states of belief where intuitively it seems that there is only one. This is because beliefs are individuated by their contents, and so if descriptive beliefs consist in being for proceeding as if p , we can differentiate between two states of being for depending on how we disambiguate “ p .”¹⁹

Thus, if there is some relation that can distinguish descriptive from non-descriptive beliefs by playing the B -role in the diagram and taking representational contents and not propositions as $relata$, it is not the relation of proceeding as if. If there is any such relation, it will be something else. Calling this putative relation pai^* , perhaps one could say in the spirit of the last response that in the descriptive case, the property $\langle pai\ p_{proposition} \rangle$ just is the property $\langle pai^*\ p_{representational\ content} \rangle$. However, this cannot be right for the simple reason that by hypothesis, propositions and representational contents are distinct kinds of entity. Therefore it is not possible that $\langle p_{proposition} \rangle = \langle p_{representational\ content} \rangle$. So it is not possible that $\langle pai\ p_{proposition} \rangle$ just is $\langle pai^*\ p_{representational\ content} \rangle$.

18 Schroeder, *Being For*, 155.

19 An anonymous referee suggests that this does not over-generate beliefs because there are two different notions of proceeding as if in play. This might be correct, but the resulting view fails to provide a unified account of belief, which was the main aim of the theory. Indeed, this seems to be a problem for the response more generally.

In conclusion, Schroeder's theory fails to achieve the main task it sets for itself: to provide a theory of belief that distinguishes between descriptive and nondescriptive beliefs. While I have rejected a number of ways of conceiving the *proceeding as if* relation for this end, I have not shown that *any* way of conceiving the relation will fail. But the onus must be on the being for proponent to show that there is some conception available.²⁰

University of Sheffield
j.ld.brown@sheffield.ac.uk

REFERENCES

- Bex-Priestley, Graham "Expressivists Should Be Reductive Naturalists." In *Oxford Studies in Metaethics*, vol. 18, edited by Russ Shafer-Landau. Oxford: Oxford University Press, forthcoming.
- Brown, James L. D. "Expressivism and Cognitive Propositions." *Journal of the American Philosophical Association* 5, no. 3 (September 2019): 371–87.
- . "Interpretative Expressivism: A Theory of Normative Belief." *Philosophical Studies* 179, no. 1 (January 2022): 1–20.
- Dreier, James. "Meta-Ethics and the Problem of Creeping Minimalism." *Philosophical Perspectives* 18, no. 1 (November 2004): 23–44.
- Gibbard, Allan. "Normative Properties." *Southern Journal of Philosophy* 41, no. 51 (March 2003): 141–157.
- . *Meaning and Normativity*. Oxford: Oxford University Press, 2012.
- Horgan, Terence and Timmons, Mark. "Cognitivist Expressivism." In *Metaethics after Moore*, edited by Terence Horgan and Mark Timmons, 255–98. Oxford: Clarendon Press, 2006.
- Köhler, Sebastian. "Expressivism, Belief, and All That." *Journal of Philosophy* 114, no.4 (April 2017): 189–207.
- Ridge, Michael. *Impassioned Belief*. Oxford: Oxford University Press, 2014.
- Schroeder, Mark. *Being For*. Oxford: Clarendon, 2008.
- . "Two Roles for Propositions: Cause for Divorce?" *Noûs* 47, no. 3 (September 2013): 409–30.

20 Many thanks to two anonymous referees of this journal for their comments. This research was supported by funding from the British Academy (grant no. PF21\210089).

NO GRIT WITHOUT FREEDOM

Berislav Marušić

GRIT IS THE TRAIT of persevering in difficult courses of action in the face of adverse odds—for example, of persisting in the pursuit of a career. In their important and interesting article, “Grit,” Jennifer Morton and Sarah Paul articulate a philosophical account of the rationality of grit.¹ The topic is important, because grit is conducive to success and flourishing, and it is interesting, because it promises a welcome enrichment of the philosophy of action by extending the focus from mundane to temporally extended, difficult action.

Morton and Paul identify grit as “a trait or capacity that consists partly in a kind of epistemic resilience,” and they defend its rationality in terms of a permissivist ethics of belief (178). The gritty agent, according to Morton and Paul, is epistemically resilient in her response to what an impartial observer might perceive as evidence of incapacity, and she is rational in doing so, insofar as such a response is permissible in her situation.

Though much of Morton and Paul’s account of grit is illuminating and plausible, I think they underestimate a crucial element that is required for the explanation of the rationality of grit: freedom. In this paper, I will explain the significance of freedom for an account of the rationality of grit and suggest that, once this is properly understood, the rationality of grit can be regarded as an instance of practical rationality.

I

Morton and Paul defend what they call the Evidential Threshold Account. They argue that “the gritty agent’s evidential threshold for updating her expectations of success will tend to be higher than the threshold an impartial observer would use” (195). This can be rational because, they suggest, as long as there is more than one rationally permissible doxastic response to a body of evidence, a “grit-friendly” epistemic policy can be defended on pragmatic grounds: “within the

1 Morton and Paul, “Grit.” Page references will be inserted parenthetically into the text.

set of epistemically permissible policies an evidential policy is better insofar as it protects to some extent against despair” (194). On Morton and Paul’s view, a gritty agent will take many setbacks not as a reason to weaken or abandon her belief in success. This is because she uses a permissible evidential policy that requires a higher evidential threshold for such belief revision.

In this explanation of the rationality of grit, freedom does not play a significant role. To bring out why this is an omission, I would like to contrast two examples: the gritty graduate student and the gritty gambler. The gritty graduate student persists in her efforts to publish a paper in a prominent journal in her field, despite repeated setbacks. Analogously, the gritty gambler persists in his efforts to win the jackpot at a slot machine. And we may suppose, not entirely unrealistically, that the odds of publishing an article in a prominent journal are similar to the odds of winning the jackpot at a slot machine. Of course, the opportunity costs will be different, though sustained pursuit of each goal will be costly, and the value of the goals will be different—so the analogy is imperfect. Nonetheless, neither the gritty graduate student nor the gritty gambler—unlike an impartial observer—takes their respective setbacks as evidence that they lack the capacity to succeed in the paths they have committed themselves to.

It is not hard to imagine the gritty graduate student to be rational. After all, this is what it usually takes in graduate school—to persevere despite considerable setbacks. In contrast, it is hard to imagine that persevering in playing the slot machines in pursuit of a jackpot could be rational. Indeed, this seems like the paradigm of irrationality. To imagine it as rational, after all, we would have to assume that the cost of playing is really low, so that the gambler is neither spending a lot of his money nor forgoing opportunities to pursue a better goal. Perhaps we have to imagine that the only real cost of playing is the time invested, so that the rationally gritty gambler would be someone with a part-time job that, though it pays no wages, gives him the chance at a one-time high payout.

But if the odds for the gritty graduate student and the gritty gambler are comparable, and if pursuit of each goal has significant costs attached to it, why does grit in one case seem paradigmatically rational, whereas in the other case it does not? I submit the following: whether one will win a jackpot at the slot machine has very little to do with one’s agency. The only involvement of the agent is the act of playing, but the agent has no influence over the outcome of the gamble. Once the coin is in the slot, the outcome is entirely determined by the machine. In contrast, if grit in the pursuit of a strong publication is rational then this is so at least partly because whether one succeeds in publishing a paper in a leading journal is *to some extent* up to the agent. Of course, obviously, it is not entirely up to the agent. However, the agent’s efforts will make a decisive

difference in whether the paper is accepted or not. One does not publish a paper in a prominent journal through sheer luck—in contrast to winning the jackpot at a slot machine.

II

One might respond that Morton and Paul can capture this observation. After all, they do say: “As we see it, the central question for an agent considering whether to persevere is, ‘Will continued effort be enough?’” (188). It seems that they do allow that what is important for an assessment of the rationality of grit is the agent’s appreciation of her effort. However, even though they recognize the significance of agency for an account of the rationality of grit, they do not take freedom to be crucial to explaining the rationality of grit but *commitment*. They write, “As a consequence of committing to a goal, the agent’s threshold should go up for how compelling new evidence must be” (194). It is through the notion of making a commitment that Morton and Paul aim to capture the thought that agency matters for understanding the rationality of grit, not through the fact that something is up to the agent.

On a first glance, this view faces two problems: first, the problem of how to coordinate assessments of evidence prior to committing with assessments of evidence afterward, and, second, the problem of how to understand the rationality of grit from the agent’s own perspective.

To see the problem of coordination, suppose that before committing to, say, a career in physics, someone judges her odds of success to be very poor. Nonetheless, for whatever reason, she subsequently commits to it. At this point, it seems that she will have to change her odds of success *just because she committed to pursuing physics*. Yet this is problematic, because it is an irrational updating procedure. If, before committing, she gave certain low odds to success, conditional on committing, then it is irrational to raise the odds *just because she made the commitment*. Moreover, after committing, she can no longer regard her earlier judgment of the odds as rational, even though she can offer no new reasons for why it was mistaken.

Morton and Paul recognize these potential problems and formulate their view so as to avoid them. They write: “the change in threshold does not apply retroactively; resolving on a goal should have no effect on how one understands the significance of the evidence one already has” (197). However, in avoiding the coordination problem, their view faces another difficulty: it turns out that grit can only be rational if, in advance of making a commitment, one has not carefully considered the evidence concerning the prospect of success in pursuing a goal. That is because the evidential threshold for assessing the odds of

success can only be as grit friendly as the evidential threshold used prior to commitment. Yet this strikes me as a flaw: it implies that the only room for grit comes from evidence that has not been considered prior to commitment. Therefore, the less consideration one has given to a project before committing to it, the more room there is to be rationally gritty afterward!

There is also a second difficulty: Morton and Paul explain the rationality of grit by appeal to permissible epistemic policies whose adoption is justified on pragmatic grounds. They are careful to distinguish their view from pragmatist accounts of doxastic rationality, according to which pragmatic grounds directly make belief rational. Instead, they opt for a tiered approach, according to which an agent's first-order deliberation is informed by exclusively evidential considerations, and pragmatic considerations kick in only at a second tier—at the justification of the agent's policy concerning how to weigh those considerations.

What is problematic for such a two-tiered approach is that it makes it hard to see how an agent could understand herself as rational in being gritty. Suppose you adopt a "grit-friendly evidential policy" and you exhibit "some degree of inertia in [your] belief about whether [you] will ultimately succeed, relative to the way in which an impartial observer would tend to update on new evidence" (194). And suppose you now meet an impartial observer—perhaps a guidance counselor or a bookie who sells bets on the outcome of your project.² You agree with them about what the evidence is, but you disagree with them about which beliefs it renders rational. You say, "I think that I will make it in physics!" The other replies, "Why will you succeed where many others have failed?"

What should you say, on Morton and Paul's view? The true answer would be that you are pragmatically justified in adopting the evidential policy you have, because you have made a commitment. However, you can neither justify your assessment of the odds by appeal to having made the commitment, nor by appeal to what justifies your use of the grit-friendly policy. That is because this justification is in the background and not something you could appeal to, at least not without falling into pragmatism. On Morton and Paul's view, "since [evidential] policies govern the way in which we respond to evidence in a given situation, they cannot themselves be called into question while first-order reasoning is in progress" (191). Yet if the standards for reasoning cannot be called into question while first-order reasoning is in progress, then it is not clear how the gritty agent can be self-consciously gritty—how she can understand herself as gritty and rational at the same time.

2 See Marušić, *Evidence and Agency*, ch. 1.3.

III

I hold that an appeal to freedom can help resolve both problems. To see this, let us return to the contrast between the agent and the impartial observer.³ I concur with Morton and Paul that they have different views of the odds concerning success in pursuit of the relevant goal. However, I do not think that this is because the agent, but not the observer, is in need of avoiding despair. Rather, it is because their relation to the achievement of that goal is fundamentally different: whether the goal is achieved depends essentially on the agent's efforts—on her exercise of her freedom. In contrast, whether the goal is achieved does not depend on the efforts of the impartial observer; it is not subject to his freedom.⁴ It is *this difference* between them that accounts for why they are rational in differently responding to the same body of evidence. Indeed, the contrast between the gritty gambler and the gritty graduate student brings this out: the less we take each of their efforts to matter, the harder it is to see the rationality of their assessment of the odds to differ from the impartial observer's. Thus, even if we can imagine the gritty gambler to be rational, we cannot imagine his odds to be any different from those of an impartial observer.

We can now hold on to the thought that grit is doxastic resilience. However, the rationality of such resilience is explained differently than Morton and Paul propose to do. What justifies the agent in responding differently to the evidence than the impartial observer is that, since it is at least to some extent up to her whether she achieves her goal, she has a different view of what is going to happen, precisely to the extent that matters are up to her. In particular, when an agent reasons about what she is going to do, her answer to that question is supposed to be settled by the very reasoning that she is engaged in, to the extent that what she is going to do is up to her. (Kant's dictum is that we act under the idea of freedom!) For the agent, insofar and to the extent that matters are up to her, the question of what she is going to do is a *practical* question.

So far, this is an observation about how an agent arrives at her initial decision about whether to commit to a goal. However, the observation can be extended to the diachronic issue of how to understand the gritty agent who displays doxastic resilience: the gritty agent persists in viewing matters as up to her, rather than undergoing a *gestalt switch* and viewing the question of her success as a simple outcome. The doxastically resilient rational agent is not (permissibly) overconfident but rather thinks about her future, insofar as it is

3 I say a little bit more about partiality in section IV.

4 Here we should assume that observation does not make a difference to the agent's actions. The case in which the agent knows herself to be observed and, for that reason, acts differently, is a special case.

up to her, in a fundamentally different way than someone who thinks merely in terms of odds. A failure of doxastic resilience is exhibited in the shift to the predictive mode—to the frame of mind in which one asks, “And what are my chances of succeeding anyway?” Indeed, it seems to me that grit is best understood in terms of a general focus on the practical—in terms of sustained attention and reflection on things one is free to *do*, rather than on things that happen to one or that are standing traits or properties of the agent.

I hasten to add that my suggestion here is not that the gritty agent ignores the evidence concerning success: doxastic resilience does not consist in an unrealistic assessment of the odds. Rather, the doxastically resilient but rational agent maintains the evidence in view, albeit not *as* evidence but, rather, as considerations of difficulty.⁵ This is so because the rationally gritty agent—unlike an inflexible or stubborn agent—is *practically* rational, and practical rationality requires a proper appreciation of the difficulty of one’s actions. Indeed, on the view I have suggested, the rationality of grit is an instance of practical rationality—of adequately responding to the practical considerations that are relevant in our context insofar as matters are up to us.⁶

Finally, although I have offered here an explanation of the rationality of doxastic resilience, I suspect that there is more to grit than such resilience. Indeed, it seems to me that doxastic resilience may be only a small, even if important, piece of the story of what grit consists in. The gritty agent does not just persist in the pursuit of a goal, despite setbacks. It would be inflexibility, not to say madness, to persist in doing the same thing only to expect a different outcome. As much as grit is about doxastic resilience, it is also a creative response to failure. The gritty agent sees setbacks as particular ways in which difficulty manifests itself and responds creatively to them, without toggling back into prediction mode. Indeed, this further brings out the significance of freedom for grit, because—as the contrast between the gritty gambler and the gritty graduate student illustrates—room for such creativity exists only to the extent that matters are up to the agent. There is no such thing as creative luck.

5 It is hard to work this out precisely. For my vexed attempt, see Marušić, *Evidence and Agency*, ch. 6.1.

6 Morton and Paul argue that “the very same exhibition of grit could count as epistemically rational in a context of privilege and epistemically irrational in a context of scarcity” (202). This strikes me as an important point. The way I would propose to capture it is that the practical situation will be different in a context of privilege and a context of scarcity.

IV

In concluding, let me now return to the two problems I discussed in criticizing Morton and Paul's view: the problem of coordination and the problem of the self-consciousness of the rationally gritty agent. On my view, what licenses doxastic resilience is not the fact that one has made a commitment, but rather the fact that something is, more or less, up to the agent and that, taking into account the difficulty of the project, it is worthwhile to pursue it. Thus, the more it is up to the agent, the more room there is for rational doxastic resilience. This means that there simply is no problem of coordinating assessments prior to and post commitment—since it is not the commitment that would license a practical view. The problem is, rather, one of understanding when things are up to the agent and to what extent. This is partly a conceptual problem, insofar as it requires a proper understanding of freedom, and partly an empirical problem, insofar as it requires a proper understanding of the facts on the ground.

As regards the self-consciousness of grit, what is crucial is maintaining a practical view. The gritty agent who is addressing an impartial observer will speak of the attractiveness of the goal she has adopted and show herself aware of the difficulty she is confronting. And perhaps, if she is philosophically sophisticated, she can point out that, as agent, she faces a practical question that the impartial observer, as observer, does not face.⁷ Ultimately, however, to bring her interlocutor to see things in *her* way, she will have to dislodge his impartiality. The other can share her assessment of her future only as someone who comes to *participate* in her pursuit of a goal, not necessarily as a joint agent, but at least as a person of trust.⁸ This suggests that it may be easier to be gritty in a supportive community—a community that shares one's outlook—rather than have to bear one's freedom alone.⁹

University of Edinburgh
 bmarusic@ed.ac.uk

- 7 This shows that, on the present account, rational belief is agent relative. However, such agent relativity should be distinguished from permissiveness: even if what it is rational to believe will be different for different agents, it need not be that several doxastic states are permissible for a single agent.
- 8 For discussion of doxastic partiality, see esp. Stroud, "Doxastic Partiality in Friendship"; and Keller, "Friendship and Belief," as well as the extensive literature that follows them.
- 9 I am grateful to Jennifer S. Marušić and John Schwenkler for comments and discussion and to an anonymous reviewer for *JESP* for helpful suggestions.

REFERENCES

- Keller, Simon. "Friendship and Belief." *Philosophical Papers* 33, no. 3 (2004): 329–51.
- Marušić, Berislav. *Evidence and Agency: Norms of Belief for Promising and Resolving*. Oxford: Oxford University Press, 2015.
- Morton, Jennifer M., and Sarah K. Paul. "Grit." *Ethics* 129, no. 2 (January 2019): 175–203.
- Stroud, Sarah. "Doxastic Partiality in Friendship." *Ethics* 116, no. 3 (April 2006): 498–524.

JOURNAL of ETHICS & SOCIAL PHILOSOPHY
<http://www.jesp.org>
ISSN 1559-3061

The *Journal of Ethics and Social Philosophy* (JESP) 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. Articles are typically published under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license, though authors can request a different Creative Commons license if one is required for funding purposes.



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.