Journal of the American Philosophical Association (2025) 211–229
DOI:10.1017/apa.2024.4

Fair Play Externalism and the Obligation to Relinquish

ABSTRACT: *This essay defends a new account of wrongful benefiting based on the principle of fair play. In particular, I argue that certain structurally-conferred group-based benefits or privileges can ground obligations on the part of innocent beneficiaries to relinquish specific gains for purposes of redistribution regardless of whether their receipt is sourced in wrongdoing or involves the imposition of harm upon relevant others. I call this approach to fair play reasoning* externalist *insofar as it turns on a novel conception of free-riding that eschews necessary appeal to beneficiaries' mental states or volition. After presenting an empirical example to help illustrate the sort of benefiting at issue and distinguishing my account from arguments rooted in the notion of structural injustice, I defend it via what I call the extension argument, respond to two salient objections, and close by suggesting its potential political utility in the American context specifically.*

KEYWORDS: fair play, free-riding, structural injustice, white privilege, wrongful benefits, beneficiary pays, moral taintedness

I. Introduction

The view that receipt of benefits, where wrongful, can call for relinquishment, even when the beneficiary themselves is innocent of any wrongdoing, is deeply intuitive and widely accepted. But what makes benefiting wrongful in the relevant sense? The two most compelling kinds of answers are when the benefits at issue are sourced in *wrongdoing*, and when their conferral involves some *harm* to others. For example, suppose a thief steals some items from V's home and deposits them at my door packaged as anonymous gifts. Proponents of what has come to be called the *beneficiary pays* principle would argue that I have an obligation to relinquish these goods once the scheme is exposed since it is wrong to benefit from others' wrongdoing or from injustice generally (Butt 2007; 2014; Haydar and Øverland 2014; Goodin and Barry 2014). The case is a bit noisy however, since my obligation here is arguably at least partly grounded in the unresolved harm to V which I am uniquely positioned to correct. To help clarify, consider a modified case in which not long after I receive the stolen goods V dies of natural causes leaving no descendants to claim compensation for the theft. While my benefiting in this case is harmless (because victimless), it's arguable that I still have an obligation to relinquish the goods once it becomes clear that my title is *morally tainted* by the thief's wrongdoing (Goodin 2013; Parr 2016). So, in fact, harm is not a necessary condition for relinquishment of benefits.

But neither is prior wrongdoing. For suppose that instead of being stolen from V and repackaged by a thief, the goods in question are deposited at my door by a hasty delivery person who simply misreads V's address. Few will deny that I still have an obligation to relinquish the goods once the mistake is revealed, and one very plausible explanation for this is the unresolved harm to V of being denied what is rightfully hers. Though here too some noise remains. To help bring this out, suppose once again that after I receive the goods, V dies before the mistake is discovered, leaving no heirs to claim compensation. While intuitions may vary here, it's arguable that my title or claim to the goods remains problematic in this case—not tainted by prior injustice, but wrongful in a sense that is capable of adding independent force to the case for relinquishment.

The foregoing discussion raises the following two general questions: Absent some connection to prior wrongdoing or harm, can innocent benefiting still be wrongful enough to ground obligations to relinquish? If so, what explains this fact? The goal of this article is to defend a new account of wrongful benefiting that is capable of helping answer these questions in relation to a specific range of cases. In particular, I will argue that the principle of fair play can call for relinquishment of certain kinds of group-based benefits or privileges, even when these are received innocently. It will be noted, of course, that none of the foregoing examples refer to *groups*, and that they all involve direct *interpersonal transactions* (i.e., some specific party confers benefits on another). It's clear however that in general wrongful benefiting can occur in a much wider range of cases, including for example, those in which benefits are conferred differentially across groups, be it deliberately (e.g., through state action) or *structurally* (see, e.g., Kim forthcoming). The present argument will focus on a particular subset of structural cases, arguing that the group-based benefiting observed there constitutes a form of free-riding that is wrongful *in itself* and can call for relinquishment of specific gains *independently* of whether they are sourced in wrongdoing or involve the imposition of harm or added burdens upon others. I call this approach to fair play reasoning *externalist* since the conception of free-riding it deploys eschews appeal to the mental states or volition of beneficiaries.

Appeals to fair play in relation to wrongful benefiting are not without precedent. For example, it has been argued that fair play can ground or help motivate obligations to: protest and resist the activities of states that wrong their own populations (Delmas 2018); respond to one's culpability as a beneficiary of collective wrongdoing (Meketa 2015), and; disgorge the benefits of public goods sourced in international state wrongdoing (Pasternak 2017). Relatedly, it has also been argued that those who benefit unintentionally from the historic emission of fossil fuels may sometimes qualify as free-riders (e.g., Gosseries 2004; Anwander 2009). Their many merits aside however, none of these arguments makes any attempt to capture the fair play principle's ability to ground obligations to relinquish *without* simultaneous appeal to facts about prior wrongdoing or occurrent harm, as I shall do here. Moreover, beyond distinguishing my argument from these other fairness-based views, it's worth noting that this reliance on *benefit alone* comes with certain theoretical advantages as well. For example, the beneficiary pays principle has been faulted for overstating the moral significance of

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

the (merely) causal connection between beneficiaries and victims of the same injustice (Parr 2016; Knight 2013; Huseby 2015). Since my view requires no reference to such a connection, this challenge cannot apply. My externalist fair play argument also shares a key advantage with the moral taintedness view in that it is less "informationally demanding" than attempts to address wrongful benefiting via compensation or restitution (Goodin 2013). Indeed, since it needn't trace the focal benefits to any prior wrongdoing, my view is arguably even *less* demanding in this sense.

But if fair play can call for the relinquishment of specific gains based on benefit alone, why hasn't this fact been acknowledged sooner? As I see it, one reason is that finding a clear empirical example of such "pure" benefiting presents certain difficulties. In the world as we know it, unfair group-based benefiting (structural or otherwise) very seldom, if ever, appears wholly disentangled from prior wrongdoing or associated harm to others. This fact may have made the issue seem immaterial or even distracting to many non-ideal theorists. A second closely related factor is that if one's argumentative goal is to convince those innocent of wrongdoing that they have nevertheless benefited wrongfully, it makes good sense to focus on benefits rooted in the most conspicuous and morally serious wrongs available, since these will be hardest to dismiss or ignore. By contrast, it's not immediately clear what practical value there could be in considering whether innocent structural benefiting alone can call for relinquishment. Even if it can, one might reasonably doubt whether that fact is likely to be what prompts innocent beneficiaries to sit up and take notice. This article challenges that presumption, insisting that beyond its theoretical value, there are important practical gains to be had by clarifying the independent wrongfulness of bare structural privilege. In particular, I will argue that a fair play analysis of contemporary inequality may be of some *political* utility in relation to the ongoing material and ideological struggle over redistributive racial policy in the United States (e.g., affirmative action, reparations, welfare).[1]

In the following section, I'll present an empirical example to help illustrate the sort of group-based structural benefiting at issue, and distinguish my approach from arguments rooted in the notion of *structural injustice*. In section III, I will introduce the idea of externalist free-riding and defend it via what I call *the extension argument*. In section IV, I respond to two arguments that my fair play account fails to advance the moral theoretic discussion around wrongful benefiting: first, that the obligations it generates are already covered by *luck egalitarianism* (e.g., Knight 2009); and second, that it continues to rely on prior injustice. Section V concludes.

II. Interpreting Inequality: A Guiding Example

In a 2020 article, behavioral economists Redzo Mujcic and Paul Frijters describe a simple field experiment in which trained volunteers from different gender and

1 My argument is not meant to turn on any particular view in the metaphysics of race, and readers are encouraged to interpret all uses of 'race' as shorthand for "membership in a racialized group."

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

racial groups boarded public buses in Queensland, Australia and attempted to pay with an empty fare card. When the scanner refused their card, volunteers told the driver that they had no money, but needed to get to a station roughly 1.2 miles away. The idea was to observe whether any significant pattern of differential treatment would emerge along demographic lines. Focusing on the axis of race in particular, it was found that over the course of more than 1500 iterations, white testers were on average twice as likely to be admitted without paying than Black testers (72% vs. 36%).

This clear racial discrepancy cries out for both explanation and moral assessment. One natural thought here is that individual drivers knowingly discriminated against Black passengers or in favor of white ones. If so, this line of reasoning goes, they should be identified and held to account. Mujcic and Frijters resist this interpretation however, noting first that they observed no evidence of "own-group bias" (with drivers of all races admitting higher numbers of whites), and second that if drivers were consciously motivated by racial hostility "then variations in the attire or displayed status of the test customers should not matter" (Mujcic and Frijters 2020: 971). When such variations were tested however, it was found that adding "positive signals of socioeconomic status and trust" (Mujcic and Frijters 2020: 972) – i.e., business suit and briefcase or military uniform – resulted in Black testers being admitted roughly as frequently as casually dressed whites (white testers in these signal conditions were admitted over 90% of the time). Based on these results and randomized follow-up interviews, Mujcic and Frijters argue that drivers' decisions about whom to admit were rooted not in racial animus but a form of "statistical reasoning" in which passengers' racial appearance (among other factors) is relied upon to make various kinds of quick inferences about their honesty, worthiness, and propensity to push back or cause trouble (Mujcic and Frijters 2020, 971–3). These spur-of-the-moment judgments tend to favor white passengers such that, over time, what may feel to drivers like discrete impartial decisions collectively give rise to the observed racial discrepancy in outcomes.

Supposing their explanation is correct, what then is the appropriate moral response? One might insist that even if drivers' discrimination wasn't deliberate, they should still be held accountable in some way. Racism of any kind is not to be tolerated, and so drivers who exhibit the relevant behavior should be identified and possibly replaced by others who display less bias. There are a number of serious limitations associated with such individualistic reasoning however. First, it wrongly implies that the issue is reducible to the actions of a few bad apples when we know that such high-speed discretionary judgments are ripe for the manifestation of widely-shared implicit racial biases. Second, such reasoning also runs the risk of obscuring how this case and others like it are symptomatic of, and entangled with, broader patterns of racial injustice.

One way to avoid these limitations is to view such cases through the lens of social structure. To help see what this entails, consider the following passage from Iris Marion Young's well-known structuralist conception of social oppression. On this account, she notes,

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

> the tyranny of a ruling group over another, as in South Africa [under Apartheid], must certainly be called oppressive. But oppression also refers to systemic constraints on groups that are not necessarily the result of the intentions of a tyrant. [. . .] In this extended structural sense oppression refers to the vast and deep injustices some groups suffer as a consequence of *often unconscious assumptions and reactions of well-meaning people in ordinary interactions, media and cultural stereotypes, and structural features of bureaucratic hierarchies and market mechanisms – in short, the normal processes of everyday life*. We cannot eliminate this structural oppression by getting rid of the rulers or making some new laws, because oppressions are systematically reproduced in major economic, political, and cultural institutions. (Young 1990, 41, my emphasis)

The core idea here is that oppression, and other kinds of structural wrongs, need not turn on any specific party's wrongdoing, nor on any particular unjust law or policy. For this reason, perceiving such wrongs often requires stepping back from a narrowly individualist or institution-level analysis and taking up a broader perspective on social affairs which prioritizes social *relations* and *processes*. We take what Young, in later work, calls "a structural point of view" on things when we "try to see how the actions of masses of people within a large number of institutions converge in their effects to produce patterns and positioning" (Young 2011, 70-1).

Although Mujcic and Frijters don't use the language of social structure, it should be clear how their account of drivers' decisions can be seen as exhibiting precisely the kind of everyday judgments and actions Young identifies as forming the substance of structural processes. From this perspective, while causally traceable to the tacitly stereotypical judgments and reasoning of particular drivers, the racial discrimination observed here *transcends* the specific individuals involved, drawing our attention to broader issues concerning the social and material significance of racial group membership in daily life. Moreover, their focus on inter-group comparisons enables us to keep grip on the ways in which this relatively modest form of contemporary racial discrimination echoes the "patterns and positioning" associated with historic practices of imposed racial hierarchy—something a more individualistic analysis may fail to register.

What kind of moral response would such a structuralist interpretation of this case call for? Throughout her work on structural wrongs, Young herself emphasized that normative judgments made from a structural point of view ask not who acted wrongly, but whether the myriad social relations and processes that constrain and guide the actions of individuals in everyday life generate "patterns and positioning" that can be characterized as unjust. Her later work specifically develops a general account of *structural injustice* as obtaining wherever socio-structural processes put some groups "under systematic threat of domination or deprivation" while simultaneously "enabl[ing] others to dominate or to have a wide range of opportunities open to them" (Young 2011, 52). The elimination of such injustice, on her view, is a matter of *political responsibility*.

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

While the details of Young's account of political responsibility are complex, for purposes of articulating the contrast with my own fair play argument, it will suffice to zero in on *five* key features of her view. First, Young insists that this species of responsibility involves no attribution of *personal culpability* on the part of its bearers. Second, rather it is *forward-looking* in that its fundamental aim is to create social conditions that are more just, rather than to rectify any specific wrongful act or event. Third, one comes to bear this responsibility in virtue of one's *active participation* within a system (local, national, global) that gives rise to the relevant structural injustices. Fourth, it follows that political responsibility on Young's view applies to *all members* of the relevant socio-structural arrangements. Lastly, while all participants bear at least *some* responsibility for eliminating structural injustice, Young maintains that *beneficiaries* of such injustice may have greater responsibilities. For example, in the closing pages of her 2011 book, Young argues that contemporary whites bear "a special moral and political responsibility to recognize our privilege, to acknowledge its continuities with historic injustice, and to act on an obligation to work on transforming the institutions that offer this privilege" (Young 2011: 187).

We are now in a position to identify some important contrasts concerning how my fair play argument will apply to the case at hand. The first thing to note is that while I too take up a structural interpretation of Mujcic and Frijters' data, and construe the wrong there as structural in nature, my account interprets that wrong as a form of *wrongful benefiting* rather than as a manifestation of structural injustice in Young's sense. The upshot of this difference in framing is significant, and can be expressed in terms of the five features of Young's view just identified. In particular, while my fair play account also implies no personal culpability on the part of obligees (feature one), it construes their obligation as a *backward-looking* response to a particular wrong, namely: externalist free-riding (contra feature two). Furthermore, while that fair play obligation is also partly grounded in obligees' participation in a political system that gives rise to structural wrongs (feature three), on my account that obligation is borne *exclusively by beneficiaries*, rather than collectively by all (contra feature four). Lastly, my account differs from Young's structural injustice argument in that it demands not only that beneficiaries "work on transforming the institutions" that afford them privilege (a quite open-ended requirement), but rather calls for the relinquishment of *specific gains received*. Taken together, these differences show that my account is meaningfully distinct from Young's framework, though not inconsistent with it.

Before we turn to the elaboration and defense of my fair play argument, it's worth pausing to acknowledge that if my goal here is to identify a guiding example in which it's clear that the obligation to relinquish is grounded in *benefit alone*, one might think the foregoing case is not ideal since the benefits enjoyed by whites arguably involve some correlative harm to non-whites. Indeed, Mujcic and Frijters' themselves describe their findings as not simply a matter of whites receiving disproportionately more discretionary benefits, but view the comparative withholding of those benefits from non-whites as a

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

nontrivial social burden.[2] Nevertheless, I believe this case represents as clear an empirical example of externalist free-riding as we are likely to find, and that it also helps to display something of the potential *political* value of my fair play account. To see what I mean by this, consider one last normative interpretation of Mujcic and Frijters' data.

According to this view, the racial discrepancy revealed in the transit study should *not* be understood as imposing any added burden upon non-whites since the benefits conferred there (being permitted to board public transit for free) are not something anyone is entitled to demand. It follows that the withholding of those benefits cannot qualify as harmful. On reflection then, there is really nothing wrongful about the study's findings after all. Some will insist (I think rightly) that this analysis is morally mistaken. Others will point out (rightly again) that it exhibits a clear failure to identify with the interests of non-whites and a marked insensitivity to broader conditions of structural racism. However, given the prevalence of such dispositions and attitudes in the American context, it would help to have an effective argumentative rejoinder to such reasoning, particularly one that is able to capture the wrongfulness of structural racial privilege without immediately invoking politically-charged and oft-dismissed concepts like systemic racism or racial oppression. Fair play externalism can be of service here, for as I'll argue in the remaining sections, even *conceding* that the racialized disparity revealed in this and other analogous cases is harmless in the sense just described, it can be shown to violate the principle of fair play and call for relinquishment on the part of beneficiaries.

III. Fair Play and Externalist Free-Riding

In general, arguments from fair play seek to ground obligations by appeal to a principle concerning the fair distribution of benefits and burdens within schemes of social or collective activity. Much of the best-known literature on fair play concerns its connection to political obligation or the obligation to obey the law (Hart 1955; Rawls 1999 [1964]; Nozick 1974; Simmons 1979). In this context, theorists have sought to determine whether the fair play principle is capable of generating obligations to participate in large-scale schemes like states by paying one's taxes, respecting the apportionment of various public goods, and so on. But the purview of that principle clearly exceeds this first-order concern, since questions of fairness do not simply end where voluntary compliance with a given scheme begins. For this reason, our discussion can sidestep ongoing controversies about the general obligation to obey the law, and follow Young in considering things in terms of political responsibility more broadly construed. The issue to be decided here is whether structural wrongs of the relevant sort can constitute violations of fair play and call for relinquishment of specific gains on the part of innocent beneficiaries who are already committed to, or at least actively

2 Estimating that around 3% of transit passengers present themselves with an empty fare card, Mujcic and Frijters calculate the value of this specific form of racial privilege at around $2.5 million annually, and the relative cost to Black persons from the lack of privilege at around $179,000 (Mujcic and Frijters 2020, 993-4).

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

participate in, the system that gives rise to those wrongs. I will argue that they can, and that appreciating this point requires recognizing those wrongs as a form of free-riding.

While the exact conditions under which one qualifies as a free-rider are disputed, the basic intuition is that it is wrong to benefit from schemes of collective activity in ways that are out of step with one's contributions thereto. Sometimes this misalignment is due to the fact that one has made no contributions at all, as in the case of someone who deliberately evades paying their transit fare, literally helping themselves to a free ride. Other times one may have made some contributions to a given scheme but still count as free-riding, as in the case of someone who fails to pull their weight in fulfilling the tasks of a group project. What seems essential is that the free-rider, as David Lyons once put it, is "an individual who tries to get (for himself) something for nothing, who tries to avoid contributing while he consumes, who tries to take advantage of the efforts and restraints, sacrifices and burdens, hardships and inconveniences of others" (Lyons 1965: 175).

Lyons' characterization also helps to illustrate the longstanding and often implicit presumption that free-riding is an *internalist* concept, that is, that free-riders act knowingly and intentionally, or more precisely, that identifying them requires reference to their mental states or volition. This is easy enough to see in the case of *voluntarist* fair play arguments which hold that an agent must willingly and knowingly accept certain scheme-based benefits in order for their failure to reciprocate to count as free-riding (Rawls 1999 [1964]; Simmons 1979). But it is no less true of *non-voluntarist* arguments which insist that it is possible to free-ride even when the focal benefits could not have been rejected (Arneson 1982; Klosko 1987a; 1987b). For even these latter arguments presume that it is not merely *receiving* benefits that makes one a free-rider, but rather the willful failure to comply with, or at least some wrongful disregard for, norms of fair play.

Closely attached to this presumed internalism is the idea that free-riding is necessarily blameworthy. One may qualify as a free-rider through negligence perhaps, but never, it would seem, without some degree of personal culpability. But is there any decisive reason for restricting the concept in this way? Might we develop a conception of free-riding which does not turn on target agents' culpable mental states? Call such an account, which defines 'free-rider' in terms of certain other non-mental facts about the party in question and their circumstances, *externalist*.

It's worth noting that an externalist account need not deny the reality or moral seriousness of free-riding that does involve knowingly taking advantage of others. It simply insists that such deliberate action is not required. Compare Ann Cudd's externalist account of *social groups* in terms of shared constraints:

> What makes a person a member of a social group is not determined by any internal states of that person, but rather by objective facts about the world, including how other persons perceive and behave toward that person. [. . .] The externalist account of social groups asserts that externally imposed constraints are necessary, and can be sufficient, for social group membership. The externalist account denies that all

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

> groups are voluntary, while allowing that voluntary actions by the members themselves can create the external constraints that compose social groups. (Cudd 2006, 37)

Likewise, an externalist conception of free-riding denies that all free-riding is volitional, while allowing that deliberate actions can create the conditions definitive of free-riding (whatever they may be). Put differently, it asserts that certain external conditions are necessary, and can be sufficient, for someone to qualify, albeit innocently, as a free-rider.

While the general idea of externalist free-riding is reasonably clear, important questions remain concerning the unspecified "external conditions" referred to here. Obviously, these will include the non-volitional conferral of benefits, but for present purposes we needn't press any further than that. For recall that the aim here is not to articulate the full range of cases in which non-volitional benefiting amounts to free-riding and calls for relinquishment. It is simply to assert this in relation to *a specific class* of externally-imposed benefits: structural group-based privilege. To accomplish this, I claim, we need only show that the wrong in those cases is sufficiently similar to that observed in paradigm cases of free-riding that the concept ought to be *extended* to apply in their case as well. Call this the extension argument. It aims to provide proof of concept for externalist free-riding while leaving open the broader question concerning its full scope.

The first step in this argument is to note that while free-riding is very often harmful, it needn't be. The fare-evading transit user imposes no added material burden on any other patron, for example. In a well-known essay on this topic, Garrett Cullity considers what it is, absent any essential appeal to harm, that makes such behavior unfair. Cullity's view is that at the most general level, free-riders enjoy a form of "objectionably preferential treatment" (Cullity 1995: 22). Assuming something like this is broadly correct, is there any principled reason to insist that such preferential treatment is only objectionable when routed through some agent's deliberate misdeed, willful refusal, or negligent omission? Put differently, why can't it be equally objectionable when conferred via structural maldistribution in the form of group-based privilege?

One potential reply here is that while it may be possible to characterize the *outcome* of such maldistribution as morally objectionable, it is inappropriate to construe the case as involving free-riding absent any culpable contribution on the part of beneficiaries. To see why this response fails, consider an analogous challenge to Young's account of structural oppression, mentioned earlier. This challenge insists that while the structural impacts she describes may be morally regrettable, even unjust, it is misleading to characterize them as *oppressive*, since oppression requires a deliberate *oppressor*. Young's response to this kind of challenge, I take it, is to insist that there is a strong case to be made for extending the scope of 'oppression' to include a broader range of cases insofar as they exhibit a similar kind of wrong, namely: large-scale group-based harms. The fact that no one is deliberately masterminding the relevant harms—that they are structural in nature—does not mean they are not rightly viewed as oppressive. A similar line of reasoning is available in the case at hand: while 'free-rider' has long

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

been used to pick out those who deliberately (or at least culpably) take advantage of collective activity, there is a strong case to be made for extending that notion to include structural cases that result in a similar kind of objectionably preferential treatment. The fact that beneficiaries in such cases don't culpably take advantage of their situation doesn't mean that they aren't well-described as free-riders.

Since this claim departs from familiar internalist accounts of free-riding, it is worth taking a closer look at the benefits received in the case under consideration in order to better appreciate the unfairness at issue. Assume that the presence of a functional transit system constitutes a public good, and that all citizens enjoy easy access to that good if they so choose, are able to pay, and follow relevant regulations. The alleged unfairness arises in connection with the fact that members of some *independently identifiable* social group come to receive *further* benefits, beyond those aimed at by the scheme, which others do not. In particular, it concerns the fact that beyond a shared ability to use the transit system, white passengers are permitted to board for free at higher rates than non-whites. My claim is that receipt of this kind of unearned socio-structural benefit or group-based privilege constitutes a form of free-riding, which can call for relinquishment of gains.

As noted in section II, some may argue that such further benefits are not wrongful since they do not involve the imposition of any obvious harm upon others. It should now be clear how a fair play analysis can (for the sake of argument) concede this point and still insist that the gains in question are wrongful when viewed through the lens of free-riding. It should also be clear where the imagined interlocutor's analysis goes awry: they assume that if all participants to a given scheme enjoy or are in a position to access the specific benefits it was designed to produce, then the provision of any *excess* benefits to some of them can be viewed as morally neutral (or perhaps even welcome). Part of what the present argument aims to show is that this is not necessarily so. Sometimes such excessive benefiting can call for relinquishment as a matter of fair play.

One might try to resist this point by asserting that norms of fair play are binding only in relation to a specific set of intentional scheme-based benefits and that discrepancies in the distribution of benefits and burdens *beyond* that set simply fall outside the principle's jurisdiction.

While there may be no decisive refutation of this minimalist interpretation of fair play, one can certainly argue that it is unduly narrow, and provide some grounds for thinking that it ought to be broadened. In particular, one can insist that when it comes to the normative jurisdiction of the fair play principle, the fundamental question is not whether the benefits at issue are part of the deliberate aims of a given scheme, but whether the resulting distribution of benefits and burdens produced by the scheme is *fair*. This is clearly a much thornier issue, subject to dispute in particular cases, but one way to bring out the broader point here is to observe that in paradigm cases of free-riding, the distributive profile of free-riders exhibits a distinctive kind of lop-sided shape owing to the fact that they benefit from collective activity without bearing the relevant costs. Notice however that this kind of distributive "distortion" can arise in other ways as well, for example, as a result of benefiting from wrongdoing. In a 2017 paper on fair play and

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

wrongful benefiting, Avia Pasternak picks up on this fact in the course of arguing that proponents of fair play have particularly strong reasons to accept the view that beneficiaries of wrongdoing have obligations to relinquish gains in compensation to victims. "If one finds this distortion in the distribution of benefits and burdens troubling in the case of cooperative schemes" she reasons, "one should find it troubling also in the case of benefiting from wrongdoing" (Pasternak 2017: 520). I am making a similar point here, but without appealing to any outside normative principle. The claim, to repeat, is simply that in our guiding example the resulting distribution is unfair, more specifically: that the whites in that case enjoy *more than their due*, and that structural privilege of this kind constitutes a form of free-riding.

Suppose it is conceded that there is some justification for expanding the category of wrongful free-riding to include certain cases of unfair structural benefiting. One might still insist that such cases must involve a particular kind of wrong in which one party *takes advantage of* or *exploits* another. In the cases under discussion however, the fact that all parties are presumed to enjoy access to the scheme's target benefits makes it hard to see who is being exploited in the relevant way. This suggests that the structural benefiting at issue shouldn't qualify as free-riding, even on this expanded conception. But recall once again that many paradigm cases of free-riding (like fare evasion) need not involve the imposition of any harm upon other parties to the scheme. With this insight in view, the challenge is easily resolved. For consider whom (or what) I take advantage of in knowingly evading my fare. The clearest answers seem to be: the other fare-paying patrons, or perhaps the scheme of public transit as a whole. Why, then, should we view these answers as inapt in relation to structural unfairnesses? While admittedly novel, it is perfectly intelligible to speak of externalist free-riders as having non-culpably *taken advantage of* those who do not receive the same structural benefits they do, or as having *exploited*, in some sense, the transit system as a whole. The fact that this form of structural advantage taking/exploitation imposes no harmful burdens on others (arguendo), doesn't show that it is not well construed as free-riding.

Picking up on this last point, a critic might insist that whatever inclination we have to view the whites in the transit study (and/or those outside the experimental context) as taking advantage of others is due to the fact that in even *attempting* to board with an empty fare card, they exhibit something akin to the culpable will of the internalist free-rider who "tries to get (for himself) something for nothing" (Lyons 1965: 175). If that's correct, it follows that there is perhaps no need for an externalist conception of free-riding after all. Notice however that while it may be true that some transit users (white or otherwise) deliberately use empty fare cards in the hopes of being waved on for free, we need not presume any such malintention for my externalist argument to go through. To see this, simply bracket out such behavior and assume for the sake of argument that all use of empty cards is rooted in non-culpable ignorance, and that in general at least, accepting a free ride out of convenience is morally permissible (and thus non-culpable). Even on these assumptions, I claim, the fact of structural privilege represents a clear violation of fair play.

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

But if we grant that accepting a free ride out of convenience is morally permissible *in general*, what is it that makes the structural benefiting wrongful in this specific range of cases? The answer is related to the fact that in those cases the relevant benefiting tracks membership in an *independently identifiable* social group. To help bring this out, consider the following counterfactual variation on Mujcic and Frijters' transit study. Suppose that instead of the findings discussed earlier, their investigation had shown that while many testers' attempts to board without paying were granted and many others' rejected, there was *no significant demographic pattern* to the results. In the end, testers of different races, genders, social statuses, etc. were all accepted and rejected at roughly equivalent rates. It seems clear that Mujcic and Frijters would have interpreted these findings as a kind of *null result*, that is, as exhibiting no social unfairness of the sort they were looking for. The whole point of their study, after all, was to see if any of the target demographic categories makes a difference in terms of how one is treated in using public transit. The intuition here is that while accepting the odd free ride as a matter of convenience is generally fine, no one is entitled to enjoy more free rides solely on account of their being white.

The independent wrongfulness of such group-based preference is a key part of what enables my view to ground obligations on the part of beneficiaries to relinquish *specific gains*, rather than the more open-ended forward-looking responsibility advanced by Young to "work on transforming the institutions" that afford them privilege. The key difference, as noted earlier, is that my view interprets things through the lens of wrongful benefiting rather than structural injustice, which places it more squarely within the domain of backward-looking corrective justice. The foregoing counterfactual variation on Mujcic and Frijters' study also helps to illustrate an important theoretical difference between my fair play account and another kind of view which might be applied here, namely: *luck egalitarianism* (e.g., Knight 2009). Roughly, luck egalitarians hold that recipients of good brute luck have obligations to redistribute unearned benefits to those with a stronger claim to them, for example, to victims of bad brute luck. To help see the contrast, note that in the imagined "null result" scenario described above, while no unfairness of the relevant sort obtains, it may yet be the case that the various passengers who were allowed to board for free can be described as enjoying good brute luck, benefiting (let us suppose) from the un-biased whims of drivers. Thus, from a luck egalitarian perspective, that random collection of transit users may still bear an obligation to relinquish gains if there are salient others with a stronger claim to them, say, fellow citizens living below the poverty line. We'll return to this point in the following section.

At this point, a critic of my view might reach for some more familiar challenges to the fair play principle, for example, to Nozick's classic objection that it allows for the arbitrary imposition of obligations upon unsuspecting parties simply by conferring benefits upon them (Nozick 1974, 90-95). Since externalist free-riding requires neither voluntary acceptance of benefits nor willful failure to comply with the fair play principle, my account may seem even more vulnerable to this worry than standard internalist ones.[3] It should be clear however that since the structural

3 This challenge and the two examples to follow are all due to (distinct) anonymous reviewers.

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

group-based benefits under discussion are not the result of any specific agent's deliberate decision, they are not arbitrary in Nozick's sense. Another crucial difference is that Nozick's examples involve no meaningful participation by beneficiaries in the relevant scheme—benefits are simply conferred on them, as it were, out of the blue. As noted above however, our focus here is on beneficiaries who are already committed to, or at least actively participate in, the broader system that gives rise to certain structural wrongs. Taken together, these differences strike me as neutralizing the force of his worries in relation to my specific fair play argument.[4]

This last point of contrast with Nozick raises an important question about the scope of my externalist account, namely, whether the benefits at issue must be tied directly to some form of deliberate localized collective activity (like the public transit scheme) in order to ground obligations to relinquish. For example, suppose it were discovered that on average whites benefit more from Samaritan assistance in times of need, specifically: that when stranded on the side of the road due to car trouble, whites are more likely to receive help (or to receive greater help) than non-whites, and that this saves them non-trivial amounts of money, not to mention time and anxiety. Would advantaged whites have an obligation of fair play to relinquish gains based on this discrepancy? I maintain that they would, and that its moral basis is essentially the same as in our guiding example. For while the absence of an explicit scheme of mutual roadside assistance may mean that no one is entitled to demand such aid, it does not immediately follow that no distribution thereof can qualify as unfair. This might be so were we to imagine the example as set within a Hobbesian state of nature, but in our contemporary context, against the backdrop of widespread participation in the shared conditions of socio-economic life, it seems entirely reasonable to view norms of fair play as operative and capable of generating obligations to relinquish unearned benefits conferred across a wide range of everyday scenarios.

This is not to say that just *any* case of innocent group-based benefiting will qualify as externalist free-riding, and we must use caution in determining the extent of the relinquishment called for in specific cases. For example, suppose it were discovered that within a given neighborhood (or city, state, etc.) whites on average had received substantially better deals in purchasing homes than non-whites (suppose for the sake of argument that prices remained within the scope of reasonable market value for non-whites). It does not follow on my account that beneficiaries of such structural unfairness must give up their homes. However, it may well follow that they are obligated as a matter of fairness to neutralize the relevant discrepancy by relinquishing the difference between the average white purchase price and the average non-white purchase price (within the relevant context) for purposes of general social redistribution.

A helpful comparison here is Jeremy Dunham and Holly Lawford-Smith's recent argument that whites may have obligations to "offset" racial privilege via financial

4 For more general discussion of how a structuralist approach to thinking about social justice relates to Nozick's entitlement theory of justice in holdings, see Young 1990 Ch. 1 (esp. pp. 28-30) and Young 2011 Ch. 2 (esp. pp. 72-74).

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

contributions to the elimination of racial difference "commensurate with the extent of one's race privilege" (Dunham and Lawford-Smith 2017: 14). Despite the superficial similarity however, their view is ultimately closer to Young's argument than to my own, falling outside the domain of corrective justice. Another crucial distinction is their claim that "an undeserved advantage conferred on grounds of race counts as race privilege only if there is corresponding undeserved *dis*advantage" and that "the question of who bears obligations to take action against racial inequality does not arise when there is advantage alone" (Dunham and Lawford-Smith 2017: 6). An upshot of my fair play account is that both of these claims are too quick: advantage alone may indeed qualify as racial privilege and call for relinquishment of specific gains if it represents a form of externalist free-riding.

Important questions remain concerning how the relevant obligations should be discharged. For example, should they be managed by beneficiaries themselves as Dunham and Lawford-Smith suggest, or captured through some public policy? Apart from noting that in principle relinquished gains needn't be understood as earmarked for any particular recipient or purpose beyond general social redistribution, I shall leave such questions aside here. For now, I hope simply to have shown that there is nothing incoherent about an externalist conception of free-riding, and that given the clear resemblance between the unfairness present in paradigm cases of internalist free-riding and in my cases of structural privilege, there is a strong case to be made for extending the notion of free-riding to include both. I hope also to have provided some defense for the idea that externalist free-riding, construed as a form of wrongful benefiting, can ground corrective obligations to relinquish specific gains for purposes of redistribution. In the next section, I'll consider two objections to my view which claim that fair play externalism fails to meaningfully advance the moral theoretic conversation around wrongful benefiting. The first argues that the obligations generated by my account are already covered by luck egalitarianism; the second argues that, despite my insistence to the contrary, my account continues to rely on prior injustice.

IV. Brute Luck and Prior Injustice

Suppose it granted that fair play is capable of grounding obligations to relinquish on the part of innocent beneficiaries in the relevant cases, even independently of facts about prior wrongdoing or occurrent harm. One might insist that such obligations are already captured by luck egalitarianism's general claim that recipients of good brute luck have obligations to redistribute those benefits to the victims of bad brute luck. This challenge asserts that while the beneficiaries in my focal cases do indeed have obligations to relinquish structural gains, this is ultimately because their receipt was unchosen and unearned. (Knight 2013, Huseby 2015, and Parr 2016 all develop a similar challenge to the beneficiary pays principle.) Indeed, one might even think that luck egalitarianism can also (already) capture the idea that these obligations are a matter of fairness insofar as it centers the fact that beneficiaries in this case have no special claim or entitlement to the structural privileges in question. If this is right, then it looks like fair play externalism—even

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

if theoretically distinct from luck egalitarianism—makes no meaningful contribution to the discussion around the redistribution of wrongful benefits.[5]

My response to this redundancy objection is to argue that while luck egalitarianism may ground obligations to relinquish unearned gains in my target cases, it does so for the wrong reason. To bring this out, consider what Cynthia Stark has called *the oppression objection* to luck egalitarianism which asserts that despite its commitment to the equal moral worth of persons, luck egalitarianism ends up condemning oppression in a merely conditional way. To explain, while luck egalitarians have found various ways to categorize oppressive social relations as unjust, they invariably explain this injustice by appeal to the fact that being subject to discrimination on the basis of one's social identity is a matter of bad brute luck. However, feminist theorists have replied that oppressive social relations are a more basic kind of wrong that is unjust *in itself* and not reducible to the fact that one's oppression is unchosen. For example, Stark points out that overlooking this point leads luck egalitarians to condone oppression when it is created by choice, as in the case of "gender hierarchy created by women's informed choices to do unpaid care work" (Stark 2020: 3). I want to make an analogous claim in relation to *privilege* of the specific sort identified here. Like oppression, this form of structural group-based benefiting (even considered independently of prior wrongdoing or occurrent harm) is an essentially inegalitarian social relation—wrongful *in itself*—which any egalitarian theory must therefore condemn unconditionally. However, to the extent that luck egalitarianism must interpret such privilege (as it does oppression) as wrongful because unchosen (i.e., because a matter of brute luck) it leaves open the possibility that some forms of group-based privilege could be permissible if grounded in the right kinds of informed choices, which I see as a normative error at least as serious as that identified by Stark.

It's also worth noting that while the redistributive obligations generated by luck egalitarianism may be a matter of fairness in some sense, the fact that they would obtain even in the "null result" scenario discussed earlier shows that it is not the specific variety of fairness at issue here. To explain, while it is arguable that both the original and counterfactual transit studies involve the unfair conferral of unchosen and unearned benefits, only the former case depicts an instance of socio-structural *privilege*. Whereas luck egalitarianism seems ill-equipped to register this distinction, my fair play argument fixes on it directly, counting that particular kind of group-based structural benefiting as a form of externalist free-riding that is wrongful in itself and not reducible to a matter of luck. This strikes me as further grounds for rejecting the claim that the obligations generated by my account are already adequately captured by luck egalitarianism.

Let us now turn to the question of whether my fair play account continues to rely on facts about prior injustice. I've claimed that existing appeals to fair play in relation to wrongful benefiting have overlooked the fact that the fair play principle can call for relinquishment on the part of innocent beneficiaries independently of appeals to prior wrongdoing or occurrent harm. The point of my guiding example was to

5 Thanks to an anonymous reviewer for pressing me to pursue this objection.

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

help illustrate this independence. But can't it be argued that since the benefits in that case are conferred structurally there is a sense in which they too are entangled with an important kind of antecedent wrong, namely, with background conditions of structural injustice? Put differently, isn't there some sense in which it can be said of the whites in that case (and analogous examples) that they have *benefited from injustice?* If so, then my approach may not move beyond the more mixed fair play arguments cited at the outset.

To strengthen this challenge, recall my claim that a structural analysis of contemporary racial inequality helps underscore its causal and normative continuity with historic practices of imposed racial hierarchy. It's no accident, for example, that in our guiding case it is the *white* passengers who are admitted at higher rates than others. This is just one of many subtle ways antecedent regimes of explicit racial hierarchy and domination continue to exert their force in the present. In the final analysis then, it arguably *is* appropriate to see advantaged parties in such cases as benefiting from injustice once we understand that the structural processes generating the relevant privileges were themselves forged through long histories of overt racial oppression.

I grant that from the perspective of concrete social history this last claim is quite correct. As a matter of moral theory however, I maintain that appeal to facts about past injustice is not required to appreciate the unfairness in my target cases or to feel the force of the corrective obligations I've argued for. We can show this via the following Twin Earth-style example. Suppose that the relevant oppressive histories had not taken place, and that Mujcic and Frijters' data had been collected in an alternative Australian context not marred by the legacies of the profound racial and ethnic injustices of decades and centuries past. Retaining the premise that the relevant racial discrepancy in treatment is rooted in socio-structural processes—now shorn of any causal connection to prior wrongdoing—would their data still depict a social unfairness calling for correction? I submit that they would. For that case displays the same objectionably preferential treatment Cullity identified as characteristic of free-riding, and which I have associated with structural privilege. This shows that while the wrongfulness of structural privilege may invariably be amplified by its causal connection with past injustice *in the world as we know it*, this connection is ultimately contingent rather than necessary.

This rather modest theoretical point, I shall now claim, is ultimately the key to appreciating what I have called the potential *political* value of fair play externalism. In particular, I believe a fair play analysis of contemporary racial inequality may be helpful in relation to the continued political struggle around redistributive racial policy in the United States specifically, for at least three reasons. First, since the normative force of externalist free-riding does not turn on facts about historic racial injustice, it follows that familiarity with the relevant history is not required to appreciate arguments deploying that concept. Even those who *are* familiar but who insist that the U.S. has entered a postracial era wherein discrimination is no longer a serious impediment to one's life outcomes should be able to agree that it is morally objectionable to receive more than one's due on account of race alone, and that such racialized benefiting may call for relinquishment as a matter of social fairness. Indeed, this point is quite consistent

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

with the resolutely color-blind framing favored by many who resist race-conscious policy as misguided, which should at least make arguments grounded in fair play harder to dismiss than some alternatives.

Second, since externalist free-riding does not presuppose any harm or deprivation to relevant others, arguments deploying that notion are well-positioned to avoid the kind of racially resentful push-back that arises around policies centering the interests or victim status of racial outgroups (e.g., Gilens 1999; Tuch and Hughes 2011) by forcing whites' moral attention back onto their own lives and circumstances. While this strategy of highlighting whites' privileged social identity by foregrounding ingroup structural benefiting comes with its own moral psychological risks (Knowles et al. 2014), these may be worth the cost. After all, the claim here is not that invoking the moral logic of fair play will immediately persuade even the most hardened racists that structurally advantaged whites have obligations to relinquish gains. The hope is that it may help at least some more centrist or open-minded whites begin to see themselves as enmeshed within racialized structures and practices *at all*, something whites as a group have long struggled to do, and which the lived experience of social dominance itself discourages (Martín 2021).

Lastly, by showing how mere social privilege can ground redistributive obligations, fair play externalism encourages the socially advantaged to reflect further on the moral implications of structural racism. For once it is acknowledged that even harmless ingroup advantage can call for correction, it's hard to see how the moral case could be weakened when similar kinds of unearned advantage are shown to come at the direct expense of racial outgroups. Such reflection may also help move some whites toward a cognitive position from which the moral force of more direct kinds of corrective racial justice discourse—often highly threatening to their moral and social identity (see, e.g., Darby and Branscombe 2014) —stands a better chance of registering.[6]

V. Conclusion

In an important early contribution to the literature on benefiting from injustice, Robert Fullinwider suggested that to be properly justified, redistributive policies like affirmative action would need to more closely resemble the workings of what he called a "game-like scheme" in which a regulatory committee "constantly monitors [. . .] and intervenes to balance off losses or gains due to infractions or violations" with a view to ensuring that each participant's outcome is "solely the result of his own unhindered efforts" (Fullinwider 1975: 313). Only such an arrangement, he claimed, would allow us to compensate those harmed by racism and sexism without imposing on the socially privileged. For while many among the latter group may have benefited from those forms of injustice, the fact that they have generally done so innocently should exempt them from any such general

6 For a defense of the focus on white structural benefiting in relation to reparations discourse specifically, drawing on recent work in empirical social psychology, see Frigault forthcoming a. For a longer argument combining that approach with the externalist account of fair play defended here, see Frigault forthcoming b.

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press

group-based imposition. The argument outlined here echoes this basic idea of making targeted interventions to correct for violations of fair play, but with a specific focus on structurally-conferred—and therefore presumptively innocent—group-based benefits enjoyed by the privileged themselves. If sound, it will have shown not only that personal innocence fails to preclude their bearing redistributive obligations as matter of social fairness, but also that such structural benefiting is wrongful *in itself*—not reducible to a matter of luck—and can warrant correction independently of its entanglement with prior wrongdoing or outgroup harm.[7]

JOSEPH FRIGAULT
CLAREMONT MCKENNA COLLEGE
joseph.frigault@claremontmckenna.edu

References

Anwander, Norbert 2009, 'World Poverty and Moral Free-Riding: The Obligations of Those Who Profit from Global Injustice,' in *Absolute Poverty and Global Justice: Empirical Data, Moral Theories, Initiatives*, ed. M. Schramm, T. Pogge and E. Mack. London: Routledge. 179–89.
Arneson, Richard 1982, 'The Principle of Fairness and Free-Rider Problems.' *Ethics* 92 (4): 616–33.
Butt, Daniel 2007, 'On Benefiting from Injustice.' *Canadian Journal of Philosophy* 37 (1): 129–52.
Butt, Daniel 2014, '"A Doctrine Quite New and Altogether Untenable": Defending the Beneficiary Pays Principle.' *The Journal of Applied Philosophy* 31 (4): 336–48.
Cudd, Ann E. 2006, *Analyzing Oppression*. Oxford: Oxford University Press.
Cullity, Garrett 1995, 'Moral Free Riding.' *Philosophy & Public Affairs* 24 (1): 3–34.
Darby and Branscombe 2014, 'Beyond the Sins of the Fathers: Responsibility for Inequality.' *Midwest Studies in Philosophy* 38 (1): 121–37.
Delmas, Candice 2018, *A Duty to Resist: When Disobedience Should Be Uncivil*. Oxford: Oxford University Press.
Dunham, Jeremy and Holly Lawford-Smith 2017, 'Offsetting Race Privilege.' *Journal of Ethics & Social Philosophy* 11 (2): 1–23.
Frigault, Joseph forthcoming a, 'Putting Whiteness First? Why Reparations Discourse Needs a Critique of White Advantage.' in *Into the Fire: The Intersection of Race and Communication*, ed. L. Harper, Vernon Press.
Frigault, Joseph forthcoming b, *Whiteness, Fair Play, and Reparation: Toward a Political Inroad*. Under contract with Palgrave MacMillan.
Fullinwider, Robert K. 1975, 'Preferential Hiring and Compensation.' *Social Theory and Practice* 3 (3): 307–20.
Gilens, Martin 1999, *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy*. Chicago: University of Chicago Press.
Goodin, Robert E. 2013, 'Disgorging the Fruits of Historical Wrongdoing.' *American Political Science Review* 107 (3): 478–91.

7 The core ideas in this essay have been in development since 2016, and shaped through conversations with too many people to name here. The earliest formulations profited from audience comments at the 2017 Pacific APA and graduate conferences at Loyola and Brown in 2016, and Princeton in 2018. Related drafts were discussed with friends and colleagues in working groups at Boston University and CU Boulder in 2019. Thanks to David Lyons, Ann Cudd, and Candice Delmas who provided helpful comments on more recent versions, and to a host of anonymous reviewers (at this and other journals along the way) whose challenges helped to greatly advance the argument.

Goodin, Robert E. and Christian Barry 2014, 'Benefiting from the Wrongdoing of Others.' *The Journal of Applied Philosophy* 31 (4): 363–76.
Gosseries, Axel 2004, 'Historical Emissions and Free-Riding.' *Ethical Perspectives* 11 (1): 36–60.
Hart, H. L. A. 1955, 'Are There Any Natural Rights?' *The Philosophical Review* 64 (2): 175–91.
Haydar, Bashar and Gerhard Øverland 2014, 'The Normative Implications of Benefiting From Injustice.' *The Journal of Applied Philosophy* 31 (4): 349–62.
Huseby, Robert 2015, 'Should the beneficiaries pay?' *Politics, Philosophy & Economics* 14 (2): 209–25.
Kim, Hochan forthcoming, 'Structural Transformation and Reparative Obligation: Reinterpreting the Beneficiary Pays Principle.' *Journal of Social Philosophy*. (Preview online.)
Klosko, George 1987a, 'The Principle of Fairness and Political Obligation.' *Ethics* 97 (2): 353–62.
Klosko, George 1987b, 'Presumptive Benefit, Fairness, and Political Obligation.' *Philosophy and Public Affairs* 16 (3): 241–59.
Knight, Carl 2009, *Luck Egalitarianism: Equality, Responsibility, and Justice*. Edinburgh University Press.
Knight, Carl 2013, 'Benefiting from Injustice and Brute Luck.' *Social Theory and Practice* 39 (4): 581–98.
Knowles, Eric D., Brian S. Lowery, Rosalind M. Chow and Miguel M. Unzueta 2014, 'Deny, Distance, or Dismantle? How White Americans Manage a Privileged Identity.' *Perspectives of Psychological Science* 9 (6): 594–609.
Lyons, David 1965, *Forms and Limits of Utilitarianism*. Oxford: Oxford University Press.
Martín, Annette 2021, 'What is White Ignorance?' *The Philosophical Quarterly* 71 (4): 864–85.
Meketa, Irina 2015, 'Honor Among (the Beneficiaries of) Thieves.' *Ethical Theory and Moral Practice* 18 (2): 385–402.
Mujcic, Redzo and Paul Frijters 2020, 'The Color of a Free Ride,' *The Economic Journal* 131 Feb: 970–99.
Nozick, Robert 1974, *Anarchy, State, and Utopia*. New York: Basic Books.
Parr, Tom 2016, 'The Moral Taintedness of Benefiting from Injustice.' *Ethical Theory and Moral Practice* 19 (4): 985–97.
Pasternak, Avia 2017, 'Fair Play and Wrongful Benefits.' *Journal of Moral Philosophy* 14 (5): 515–34.
Rawls, John 1999 [1964], 'Legal Obligation and the Duty of Fair Play.' in *Collected Papers*, ed. Samuel Freeman. Cambridge: Harvard University Press. 117–29.
Simmons, A. John 1979, *Moral Principles and Political Obligations*. Princeton: Princeton University Press.
Stark, Cynthia A. 2020, 'Why Luck Egalitarianism Fails in Condemning Oppression.' *Feminist Philosophy Quarterly* 6 (4): 1–17.
Tuch, Steven A. and Michael Hughes 2011, 'Whites' Racial Policy Attitudes in the Twenty-First Century: The Continuing Significance of Racial Resentment.' *Annals of the American Academy of Political and Social Science* 634 (1): 134–52.
Young, Iris Marion 1990, *Justice and the Politics of Difference*. Princeton: Princeton University Press.
Young, Iris Marion 2011, *Responsibility For Justice*. Oxford: Oxford University Press.

https://doi.org/10.1017/apa.2024.4 Published online by Cambridge University Press