Journal of the American Philosophical Association (2026) 217–234 © The Author(s), 2025. Published by Cambridge University Press on behalf of the American Philosophical Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.
DOI: 10.1017/apa.2025.13

What If Everyone Did That?

ABSTRACT: *Orthodox contractualists and rule consequentialists think that, for any action, the consequences of everyone performing that action determine whether that action is permissible. For them, "what if everyone did that?" is the fundamental moral question. By making "what if everyone did that?" the fundamental question of good moral reasoning, these moral theories can easily justify prohibitions on free-riding. But it also makes them face the ideal world problem. I argue that it was a mistake for moral theorists to generalize from an objection appropriate to cases of free-riding to all of morality. In short, we should understand the question "what if everyone did that?" to give expression to one,* and only one*, kind of objection to one's action—namely, that, by performing that action, one would be making an exception of oneself. If we limit the scope of "what if everyone did that?" in this way, we can justify prohibitions on free-riding while avoiding the ideal world problem.*

KEYWORDS: consequentialism, contractualism, fairness, free-rider problem, ideal world problem

1. Introduction to Our Puzzle

Suppose that we are walking around our neighborhood, and I am about to throw some trash onto the sidewalk, knowing that most other people hold on to their trash until they get home. To stop me, you might ask me "what if everyone did that?" In asking me this, you are trying to get me to see that, by littering, I would be doing what philosophers and social scientists often describe as free-riding.

The idea of free-riding on the contribution of others to some important public good has played a central role in 20th century moral and political theory. For each of what Hume (2011) once called the "artificial virtues"—the keeping of promises, respect for the property rights of others, and obedience to the law—there has been some influential analysis in terms of the wrong of free-riding. Perhaps it is true, as Rawls (1999) argued, that we must keep our promises and respect the property rights of others because our failure to do so would amount to free-riding on the practice of agreement-making, on the one hand, and ownership, on the other. And, to take one more example, it may be true, as Klosko (1992) argued, that breaking the law is wrong because it involves enjoying the public goods made possible by a coercive state —goods such as national defense, public infrastructure, and public education— without constraining one's behavior in the ways that are necessary for the state to provide them. Nonetheless, if we are without a satisfactory justification for why we should not free-ride, we will be no closer to possessing a firm grounding for the

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

artificial virtues, not to mention why my littering in our neighborhood would be wrong.

For this, we shall need to turn to theories of what it is for an action to be permissible, forbidden, or required. Act-based moral theories—such as act consequentialism and act contractualism—have been thought to struggle with accounting for prohibitions on free-riding. Whether we ought to do whatever would result in the most well-being (Kagan 1989; Portmore 2011) or whether we ought to do whatever would minimize the strongest complaint (Sheinman 2011; Bourguignon forthcoming), it is not clear how my single act of littering could be wrong. Given that most others hold on to their trash until they get home, my single act of littering will not prevent our neighborhood from being a pleasant place to walk around. And, even if many in the area did litter, my holding on to my trash in this instance could not make our neighborhood pleasant when it would not be otherwise.

One influential strategy for defending prohibitions on free-riding involves consequentialist and contractualist moral theories taking the content of your reproach very seriously—by, roughly, making the permissibility of an action depend on the effects of everyone performing it. To adopt this strategy would be to accept rule consequentialism or orthodox contractualism, which determine what is required by seeing which principles, if generally accepted, would result in the most well-being or would minimize the strongest complaint (Hooker 2000; Scanlon 1998). Here is how the adoption of one of these views would solve the free-rider problem: If everyone in our neighborhood littered instead of holding on to their trash until they got home, our neighborhood would have trash everywhere and would be an unpleasant place to be. By contrast, if everyone in our neighborhood held on to their trash until they got home, our neighborhood would be a clean and enjoyable place to walk around.

The move to rule consequentialism or orthodox contractualism, however, raises problems of its own. According to the ideal world problem, sometimes these theories recommend acting in accordance with a principle that would be great if it were generally accepted but a nightmare to follow in situations where it is not. For example, just because a perfectly pacifistic world would be a lovely, peaceful place to live does not mean that, in our world, where there are violent people, we can never use violence in defense of ourselves or innocent others. How, then, can a moral theory condemn my free-riding without having the very property that makes moral theories face the ideal world problem—namely, the focus on the question "what if everyone did that?" What we have, here, is a puzzle.

When I was considering littering and you asked me "what if everyone did that?", you probably thought that you were raising a very particular kind of moral objection, but not the only kind. One way of understanding rule consequentialism and orthodox contractualism is that they would disagree with you. In thinking that, for any action, what would happen if everyone performed that action determines whether that action is permissible, they take "what if everyone did that?" to be the fundamental moral question. My main aim is to argue that this is a mistake. In contexts of the relevant kinds, we should understand the question "what if everyone did that?" to give expression to one, *and only one*, kind of objection to one's action—namely, that, by performing that action, one would be making an exception of

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

oneself. A moral theory that interprets "what if everyone did that?" as having this more limited role in good moral reasoning can both justify prohibitions on free-riding and avoid the ideal world problem. And act-based theories, in particular, can interpret "what if everyone did that?" in this way by taking the question either to give expression to a strong individual complaint against my littering or to give expression to the fact that, by littering, I would be making things go badly for others by making an exception of myself.

Discussion proceeds as follows. Section 2 discusses why "collectivist" views—such as rule consequentialism and orthodox contractualism—seem better placed to derive prohibitions on free-riding than their act-based counterparts. Section 3 argues that the very feature of collectivist views that makes them well-placed to derive prohibitions against free-riding makes them face the ideal world problem. This produces our puzzle: Find a moral theory that can generate prohibitions on free-riding without also raising the ideal world problem. Section 4 defends a solution on which act-based moral theories allot to "what if everyone did that?" the role that it has in commonsense moral thought—as merely one sort of moral objection that can be appealed to amongst the many others in one's répertoire. In particular, "what if everyone did that?" gives expression to the objection that one would be making an exception of oneself by performing the action in question. Section 5 refines my solution to our puzzle in light of a challenge from Brennan and Lomasky which purports to show that "what if everyone did that?" is too coarse-grained a question to distinguish between fair and unfair actions. Section 6 distinguishes my account of what it is to make an exception of oneself from some other related accounts in the literature.

2. The Free-Rider Problem

Act-based moral theories trace the permissibility of one's action to the effects, whether good or bad, of that action in particular. Consider first what is perhaps the most well-known act-based moral theory, namely act consequentialism. According to this view, you ought to always act in whatever way, of those available to you, would make things go best (Kagan 1989; Portmore 2011). The permissibility of one's action, then, on this view, depends on what difference you would make in the world, good or bad, as a result of your performing it.

What is more, the contractualist variant of this view—act contractualism—is no different on this score. According to this theory, you ought to perform, of those actions available to you, whichever one would minimize the strongest complaint (Sheinman 2011; Bourguignon forthcoming). This view departs from act consequentialism by not tracing the permissibility of your action to the *sum* of well-being it produces. Instead, act contractualism only attends, when deriving duties, to the relative strength of complaints that can be lodged by individual people, one by one. But this difference does not militate against the fact that act contractualism shares with act consequentialism a common structure: what you ought to do depends on what you can bring about by acting.

These moral theories, moreover, appear to struggle with deriving prohibitions on free-riding precisely because they have this structural feature. Typically, the cases

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

governed by prohibitions on free-riding are "collective harm cases"—cases where enough of us doing something would cause harm or the destruction of a public good but any one of us doing that thing would not cause any harm considered on its own (Nefsky 2019). Think back to our littering case. It takes precisely this form. Enough of us littering would result in the neighborhood being an unpleasant place to live. But any one act of littering on its own does not seem to make a difference to the overall cleanliness of the neighborhood.

If, however, my act of littering would not on its own make any difference to the pleasantness of our neighborhood, then how could either of the act-based theories we have mentioned condemn it? Take act consequentialism first. Either I free-ride or I do not. If I do free-ride, then there is still the public good, and I am happy for not having to restrict my liberty. If I do not free-ride, then there is the same public good, but I am less happy for having to restrict my liberty. This means that more aggregate good would be produced by my free-riding than by my contributing. So, act consequentialism recommends that I free-ride. And similarly for act contractualism. If my littering would not make a difference to the security of the public good in question, on what grounds could one lodge a complaint against my decision to litter? Indeed, it seems like *I* would have a complaint against being prevented from littering. By being prevented from littering, I am being forced to restrict my liberty for no good reason. Either way, no one in the neighborhood will be any better or worse off.

Perhaps, then, a more "collectivist" form of moral reasoning is needed to get a complaint against my littering into view. Rule consequentialism and orthodox contractualism—both of which trace the permissibility of an action to the effects of *groups* of people's actions—supply promising strategies for deriving prohibitions against free-riding. Take rule consequentialism, first. According to this theory, we ought to do whatever would be prescribed by a principle whose general acceptance would make things go just as well as any alternative principle. Now consider two principles: one that would permit people to free-ride whenever their doing so would not make a difference to the public good in question and one that forbade people from free-riding even in those cases in which one could free-ride without detracting from the public good. The rule consequentialist question, then, becomes: Where is there more aggregate goodness? The world in which a principle that permits people to free-ride when doing so makes no difference is generally accepted? Or the world in which a prohibition on free-riding even when doing so makes no difference is generally accepted? It seems like there is more overall good in the *latter* world—the world in which free-riding is prohibited. For, in the former world, there will be no more clean streets. If everyone governed their practical reasoning in terms of a principle that permitted them to free-ride when their individually doing so makes no difference, everyone would throw their trash on to the ground instead of holding on to it until they return home. Even though each of these people's individual acts of littering would not make a difference considered *one by one*, all of their individual acts of littering *taken together* would. By contrast, the only issue in the latter world will be that people have to restrain their liberty. But that is less bad overall than the loss of an important public good would be. The same move, moreover, could be made by orthodox contractualists. Surely, the complaint that an important public

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

good is gone is weightier than the complaint that one has to restrict their liberty in pursuit with others of something that will benefit all of them.

The lesson, here, seems to be that when it comes to generating prohibitions on free-riding, it is not enough to have a moral theory that makes the permissibility of your action depend *entirely* on what you can bring about yourself by performing that particular action. This is because paradigm instances of free-riding do not make any difference to the existence and quality of the public good in question. But a "collectivist strategy" seems more promising in collective harm cases. If what you ought to do depends entirely on what would happen if everyone did what you did, then the fact that everyone's littering will make the neighborhood unpleasant can make your free-riding impermissible.

3. The Ideal World Problem

As mentioned in the introduction, however, the move toward a more "collectivist" view comes with a hefty cost. In particular, rule consequentialism and orthodox contractualism face the ideal world problem precisely because they determine what you ought to do by looking at the effects of groups of people acting in the way that you are considering doing. According to the ideal world problem, sometimes collectivist moral theories recommend complying with a principle that would be great if it were generally accepted but a nightmare to follow in situations where it is not. Consider, for example, a principle called "the rule of pacifism (ROP)," according to which no one is ever permitted to use violence. If ROP were generally accepted, then the world would be a lovely, peaceful place to live because there would not be any violence. So, ROP cannot be outranked—on either rule consequentialist or orthodox contractualist grounds—by any alternative principle. So, according to both collectivist views, no one is ever permitted to use violence. As a matter of fact, however, not everyone has internalized ROP. There are violent, homicidal people among us. So, if we complied with ROP, then we would not use violence against them to defend ourselves or others. This would be a nightmare (Parfit 2011: 314).

Perhaps a collectivist will respond by putting forward what they might call the "conditional rule of pacifism (CROP)," according to which no one is ever permitted to use violence, unless someone is violent first, in which case it is permissible to use violence in self- or other-defense (provided that violence is necessary and proportional). Then, the collectivist might press: Would a world in which everyone accepts CROP not be similarly peaceful and lovely? After all, no one would ever use violence (even in self- or other-defense) in a world where everyone accepted this principle, since no one would feel free to use violence first. The problem with this move is that it fails to establish what the collectivist needs to establish to avoid the ideal world problem—namely, that CROP is better than ROP. The worlds in which each respective principle is generally accepted are identical in all the relevant respects. Even worse, "going conditional" does not even allow the collectivist to rank CROP higher than the following, absurd principle about when violence is permissible: Never use violence, unless someone is violent first, in which case kill as many people as you can. Just like with CROP and ROP, a world that was just like

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

ours except that this absurd rule is generally accepted is one where no one uses violence at all (Parfit 2011: 316-17; Rumbold 2024: 4).

A collectivist might instead respond by invoking a partial acceptance version of his view, fiddling with the *degree* of acceptance of a principle that matters for determining whether an action is permissible. A consequentialist example of this possible collectivist move is provided by fixed-rate rule consequentialism, according to which what is morally permissible is determined by those principles acceptance of which by 90% of everyone would result in more well-being than were 90% of everyone to accept any alternative principle (Hooker 2000: 80-5). And it would be trivial to formulate an orthodox contractualist version of this move. Here is how adopting a fixed-rate collectivist theory is supposed to help deliver the right result that ROP is not as highly ranked as any alternative principle: If only 90% of everyone accepts ROP, then some people would be violent and those others who *do* accept ROP would not feel free to use violence against the former group even when it is necessary to defend themselves as well as proportional to the violence (or threat of it) that they are victims of. This would result in innocent people being prevented from defending themselves against aggressors—which would involve less well-being and raise stronger individual complaints than CROP.

This collectivist response, however, also misses the mark. Fixed-rate collectivist views deliver the wrong result about other, more complex cases which also raise the ideal world problem. For example, suppose that if the principle "give to poverty relief if you are not impoverished yourself" were accepted by 90% of everyone, then global poverty would be alleviated. Suppose further that this principle is not accepted by 90% of everyone, and as a result, poorly funded charities do more harm than good with the money that is given to them. We can imagine that this is because those who run these poorly funded charities have responded, and will continue to respond, quite vindictively to the fact that it is not customary for the financially comfortable to give to poverty relief. Call this possible world "vindictive world" (Salomon 2024).

Intuitively, those who are financially comfortable in vindictive world should not give to poverty relief, since doing so will result in more harm than good. But fixed-rate collectivist views cannot deliver this result. Things would go better in vindictive world if 90% of everyone who is financially comfortable were to give to poverty relief than it would if 90% of everyone who is financially comfortable felt free not to give. This is because, by hypothesis, if 90% of those well-off were to give, then global poverty would be solved—a great thing, indeed. Moreover, vindictive world would present a problem not only for fixed-rate collectivist views but also *for any other partial compliance view*. In order to make the case of vindictive world pose a problem for a partial compliance collectivist view, one need only do the following: First, determine the rate of acceptance R at which the target partial compliance view thinks that candidate moral principles should be assessed. Next, alter the details of vindictive world such that charities will do more harm than good with gifts given to them whenever the principle "give to poverty relief if you are not impoverished yourself" is accepted at a rate lower than R. What this shows is that the problems that collectivist views face by assigning the question of "what if everyone did that?" such a central role in their theory cannot be solved by making fine distinctions among just how many people should be imagined doing something when determining whether it

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

would be permissible to do that thing (Podgorski 2018). The ideal world problem cuts deeper than that—calling into question the very idea of a moral theory which makes the question "what if everyone did that?" a cure-all.

4. My Solution

By (roughly) making the effects of everyone's doing something both necessary and sufficient for determining what we ought to do, collectivist views assign the question "what if everyone did that?" a particularly fundamental role to play in good moral reasoning. This allows them to solve the free-rider problem, but it also makes them face the ideal world problem. How do we escape this dilemma? As I have already alluded, we can do so by taking "what if everyone did that?" to only give expression to a particular kind of objection that one could make to an action. In particular, we should understand "what if everyone did that?" as a question you can ask me to get me to see that by doing what I intend to do I would make an exception of myself and, so, do something unfair. By doing this, we can solve the free-rider problem and avoid the ideal world problem, all the while not departing from an act-based structure, where the permissibility of an action is traced to which outcome that action would bring about.

4.1. Unfairness as Part of the Broad Outcome of an Action

In order to see how an act-based moral theory has the resources to make use of the fact that an action is unfair in determining the permissibility of that action, we will need to make a distinction between two notions of the outcome of an action: narrow and broad. The narrow outcome of an action is exhausted by the sum of that action's downstream effects. The broad outcome of an action, on the other hand, includes, in addition to that action's downstream effects, the act itself (Munoz 2021). This distinction between two kinds of outcome is familiar from the "consequentializing" program—where it is employed to argue that act consequentialism, despite initial appearances, can make sense of constraints on action (such as, e.g., the fact that a surgeon is forbidden from killing her healthy patient to save five others). According to consequentializers, once we are comparing the broad outcomes of the surgeon's two available actions—namely, killing her healthy patient or not doing so—we can see that killing her patient would bring about a worse outcome than not doing so. Here is their thought: If we are comparing the broad outcomes of the surgeon's killing and not killing, then we will include in this comparison the badness of killing itself. In addition, it is plausible that the action of killing one person itself is worse than letting five others die. So, in order to bring about the best outcome, the surgeon must let the five others die (Suikkanen 2020: 26).

Quite a bit of ink has been spilled over whether this consequentializing program will ultimately prove successful (Portmore 2009). Fortunately, for the purposes of our discussion, we need not delve into this debate. We need only make the observation that, when assessing the permissibility of an action, act consequentialists possess the resources to include the goodness or badness of that action itself as well as that of the other options available to the agent. And the same is true for act contractualists.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

When assessing individual complaints against an action in the moral reasoning distinctive of their view, act contractualists can consider the goodness or badness of the act itself in addition to its downstream consequences.

Notice, now, that the argument discussed above for why act-based views struggle with accounting for the wrongness of free-riding assumed that act-based views are only concerned with the narrow outcomes of actions. This can be seen by remembering that such an argument relied on pointing out that my act of free-riding, considered on its own, would make no difference to whether our neighborhood was a pleasant place to walk around. The objection to act-based views, that is, relied on my single act of littering not having any *downstream causal consequences* for whether our neighborhood was relatively clean. If act-based views can only consider the narrow outcomes of actions, when determining their permissibility, then this argument goes through without incident.

But, as we now know, act-based views can consider *more* than just the narrow outcomes of actions. They can assess broad outcomes, too, when determining the permissibility of an action. And this is how act-based views can make use of the idea that "what if everyone did that?" gives expression to the objection that one is making an exception of oneself. My littering, considered on its own, is unfair. And its unfairness can either serve to make the outcome in which I litter worse by consequentialist lights than the outcome in which I hold onto my trash or it can provide one with a strong objection to my littering in contractualist moral reasoning.

4.2. Free-Riding as Unfair

Let me first defend the claim that my littering is unfair because it involves me making an exception of myself before I elaborate on how act-based views can harness this observation in their respective models of good moral reasoning. It seems plausible that making an exception of oneself involves treating oneself as special in an objectionable way. When there is some important good in existence that clearly required the restraint of others to produce, one treats oneself as special by thinking that one need not contribute oneself. And this is exactly what I do when I litter. I know that the pleasantness of our neighborhood's streets required others to inconvenience themselves by holding on to their trash. But I take myself to be free to litter anyway. So, I can be accurately described as favoring myself over others in my neighborhood given some intuitive understanding of that phrase. What I express, then, when I litter, is a kind of partiality to myself that is morally objectionable.[1] And that is why you try to get me not to litter (or scold me after the fact) by asking me "what if everyone did that?" By pointing out to me that everyone else's failure to hold

[1] The idea that free-riding is wrong because it evinces an objectionable form of partiality is hardly unfamiliar. Nice discussions of this idea can be found, for example, in Cullity (2008), Trifan (2020), and Wieland (2024). The main contribution of this article lies not in its tracing the wrongness of free-riding to the wrong of making an exception of oneself. Rather, my main aim is to establish that the wrongness of making an exception of oneself is merely one kind of moral wrong among many rather than the fundamental wrong of which all other moral wrongs are instances. It is my contention that collectivist views face the ideal world problem precisely because they generalize from an objection appropriate to cases of free-riding to all of morality.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

onto their trash would lead to disgusting streets, you are trying to get me to see that I would be favoring myself by littering. In particular, you hope to get me to see that, by failing to hold onto my trash, I would be depending on others to keep our neighborhood pleasant.

But, then, if my littering would be unfair, then *my littering's being unfair* would be part of the broad outcome of my action of littering. Unfairness is bad, and it is bad *for* those treated unfairly— for those, in this case, whose interests are treated as mattering less than mine and for no good reason. And, since unfairness is bad for those whom I arbitrarily favor myself over, the unfairness of my littering detracts from the sum of well-being that my littering would produce. It is plausible, moreover, that the state of affairs where I behave unfairly but get to litter if I want is worse than the state of affairs in which I do not make an exception of myself but must hold onto my trash until I get home. So, an act consequentialist can infer, it would be impermissible for me to litter.

An act contractualist can make a similar move. If complaints can be made in act contractualist reasoning on the basis of the broad outcome of my littering, then a complaint can be made by an individual in my neighborhood against my littering on the grounds that my littering is unfair. And the complaint is this: You are treating me unfairly by arbitrarily favoring yourself and your own interests over me and mine. Plausibly, moreover, the complaint that my action is unfair is stronger than my complaint that I have to hold onto my trash until I get home. Since act contractualists think we always ought to do what will minimize the strongest complaint, they will deliver the intuitive result that I ought not to litter.

What is more, other non-collectivist views of good moral reasoning can employ these fairness-based complaints, too, to derive prohibitions on free-riding. For example, maxim-based views determine what I ought to do by asking (roughly speaking) whether my doing that thing repeatedly over time would minimize the strongest complaint or, on the consequentialist variant, bring about more aggregate well-being than if I were to do something else repeatedly over time (Salomon 2024). Much like act-based views, maxim-based views might first appear to be unable to derive prohibitions on free-riding on their own. After all, so long as most others continue to hold on to their trash until they get home, not even my littering over time would prevent the neighborhood from being pleasant. *Some* attention, then, to the effects of everyone's doing something seems requisite for maxim-based views' capacity to justify duties not to free-ride, too. That attention can be paid by recognizing that "what if everyone acted in the way you are acting over time?" is an objection that can be lodged against a free-rider's maxim or policy—an objection that gives expression to the fact that free-riding whenever doing so would not make a difference involves making an exception of oneself or adopting an unfair policy.

It is important to note, here, that such complaints of unfairness would be inappropriate to lodge against the well-off people who do not donate in vindictive world. This is because a well-off person who does not donate in vindictive world would not be *depending* on other people to alleviate global poverty, since, by hypothesis, they themselves would not be donating either. The same thing, moreover, could be said about a different sort of collective harm case than the one which has centrally occupied our discussion—namely, a case where no one has begun

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

to take steps to lower their carbon footprint by taking public transportation instead of driving gas-guzzling cars to and from work. By driving my gas-guzzler to and from work, I would not be depending on others to mitigate the harms of climate change. None of *them* are taking steps to lower their carbon footprint either. A more general lesson emerges from these cases concerning which sorts of effects of everyone's doing something we ought to be concerned with when determining whether someone would make an exception of themselves by doing that thing. We ought to be concerned in particular with disaster *striking* or a public good *going away* when we consider the effects of everyone's doing something, since it is only by looking at those sorts of effects that we can draw someone's attention to the fact that they would be depending on others to do something in a way that evinces their mistaken opinion that they are somehow special from the perspective of morality's demands.[2]

4.3. The Defeasibility of Unfairness

We have just seen how the fact that an action is unfair can constitute a strong objection to its performance and that pointing out the disastrous effects of everyone's performing that action would be an appropriate way of lodging such an objection. Nonetheless, it is worth keeping in mind that the fact that an action is unfair does not necessarily constitute a *decisive* objection to its performance. To see this, suppose that you and I live in a fishing community that uses traditional small nets to capture fish locally, and everyone captures just enough fish for themselves. If we all begin to use modern ships and large nets, fish stocks will be depleted. But, before others do so, I could use a larger trawler and a bigger net. There would be enough fish, and I could use the extra income to help the community. Some have thought that, in cases like this, it would be permissible for me to use a larger trawler and a bigger net.[3] But, of course, my using this more modern technology would count as unfair, on the understanding of "unfairness" defended thus far in this article. For, by hypothesis, if everyone in our community were to start using modern ships and large nets, then fish stocks would be depleted and, thus, an important public good would be destroyed.

I am not as sure as some that it would be permissible for me to use the bigger boat. Nonetheless, my account has the resources to make sense both of the thought that my doing so would be permissible and the thought that my doing so would not. To see how, first notice that it *does* seem plausible that my using the modern fishing technology unilaterally would be unfair. It seems like you would be well within your rights to ask me: "Why do *you* get to be the one who is allowed to use the modern fishing technology rather than me or one of our other fellow fishermen? Each

[2] This, it seems to me, is the truth behind Parfit's claim (2011: 305-6)—echoed by Murphy (2021: 263)—that the question "what if everyone did that?" only makes sense to ask when people are *not* in general doing that.

[3] I borrow this case from Suikkanen (2024). In his article, Suikkanen adopts an alternative, but related, strategy for accounting for collective harm cases without raising the ideal world problem. Although Suikkanen's "hybrid" version of consequentialism does not appeal to fairness or, more specifically, exception-making, it is similar to mine in that it says that the effects of everyone's performing some action, even if not dispositive, are relevant to the permissibility of that action.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

one of us would like to be the one who gets to use the bigger boats and nets, but none of us use them because doing so would involve singling ourselves out for special treatment."

Nonetheless, if unfairness is either, in a consequentialist idiom, one bad among others to be summed up along with the goods occurrent in a state of affairs or, in a contractualist idiom, one source of complaint among other possible sources of complaint to be compared pairwise, then we should expect that sometimes the fact that an action is unfair does not guarantee that it is impermissible. Sometimes, the fact that an action is unfair can be *outweighed* by other facts about the situation in which the agent might find his or herself. Perhaps this is what is going on in the fishing case. It may be that the thing to say about my using a larger trawler and a bigger net is that it is permissible, because the fact that I could bring about so much good for the community by doing so *outweighs* the fact that doing so would be unfair.[4] I need not take a stand on this question of substance, here. I will just note that whether or not you think my using a large boat would be permissible must turn on whether the fact that I could do good things for the community by doing so outweighs the fact that my doing so would be unfair.

4.4. Summarizing the View

I have been arguing that act-based approaches to moral theorizing can use the question "what if everyone did that?" to capture the idea that free-riding is wrong *because it is unfair*. This use of the question allots it a quite limited role in the moral reasoning distinctive of their views, and it is this question's limited role that allows these theories to avoid the ideal world problem. This is the punchline: just because act contractualism and act consequentialism can make use of the question "what if everyone did that?" does not mean that they face the ideal world problem. According to the act-based moral theories being defended here, facts about what would happen if everyone performed the action under consideration are relevant to whether that act is unfair. Nonetheless, these theories do *not* take the fact that everyone's doing something would be lovely to be sufficient to place us under a moral requirement to do that thing. On the contrary, act contractualism and act consequentialism can vindicate our judgment that violence in self- or innocent other-defense is sometimes permissible. In a world where some people are violent, my using violence in self-defense could result in the most overall well-being or minimize the strongest complaint. (Surely, for example, my attacker's complaint against being harmed is weaker than my complaint against being harmed without provocation.)

[4] Of course, if act contractualists want to make use of this line of thought, they would need to identify a complaint *that a single individual has* against my not using the modern fishing technology that is strong enough to outweigh the complaint that I would make an exception of myself by using it. That is because, like orthodox contractualists, act contractualists are committed to the individualist restriction on what kinds of complaints can be lodged in their moral reasoning.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

To summarize, here is the picture. An action is unfair (in the sense of involving the actor making an exception of his or her self) when and because, if everyone were to perform that action, disaster would strike or an important public good would go away. This claim about what it is to make an exception of oneself provides us with the following test for whether one's performance of an action would involve one making an exception of oneself: If one acts in such a way that not everyone else could act without causing disaster or the destruction of an important public good, then one makes an exception of oneself by acting in this way. I will call this test for exception-making the "what if everyone did that?" test and the use of this test "what if everyone did that?" reasoning.

5. Brennan and Lomasky's Challenge

As I noted in the introduction, however, the "what if everyone did that?" test should only be adopted as a first pass. This is because, as currently stated, it is vulnerable to a challenge put forward by Brennan and Lomasky (2000: 75-9). The particular challenge of Brennan and Lomasky that I have in mind targets a popular argument for the duty to vote, which goes like this: "But what if everyone were to stay home and not vote? The results would be disastrous! Therefore, I (you/she) should vote" (2000: 75). This argument from generalization is (roughly) one way of expressing my "what if everyone did that?" test. And it seems promising because, like littering, voting has a collective harm structure. Enough of us staying home instead of going to the polls would undermine an otherwise well-functioning democracy. But any particular person staying home does not make a difference.

Their challenge to the argument from generalization begins from a commonsense verdict about the following case:

> [Suppose] that Dalrymple is considering leaving the farm to pursue a career as a dentist. The question is put to her, "What if no one grew fruits and vegetables?" The result, Dalrymple admits, would be disastrous. Therefore, it is urged, she would do wrong to abandon farming for dentistry. Against this suffices the retort, "But *not everyone will* give up farming" (2000: 76)!

It is supposed to be clear, here, that it would not be unfair for Dalrymple to leave farming for dentistry just because, were everyone in her community to do so, disaster would strike in the form of something like a famine. But, Brennan and Lomasky continue, the argument from generalization predicts that we should conclude that Dalrymple would be wrong to leave farming for dentistry. So, they conclude, the argument from generalization is invalid and, accordingly, cannot ground a duty to vote. Or, in our parlance, they conclude that "what if everyone did that?" cannot provide a satisfactory test for whether one would make an exception of oneself.

They do, however, think that *sometimes* doing something that not everyone could do without causing disaster is unfair. In order to demonstrate this, they have us consider another case:

> Throckmorton is about to take a shortcut across the newly seeded lawn and is brought up by the reproof, "Well, what if everyone were to walk

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

> across the lawn? All the grass would be killed!" The reply, "But not everyone will cut across the lawn; most people take more heed of signs than I do," carries distinctly less conviction (2000: 76).

According to Brennan and Lomasky, then, the fact that, if everyone were to do what Throckmorton plans to do, an important public good would go away *does* imply that Throckmorton would act unfairly by taking a shortcut. Their question becomes "why?" What explains why we have different judgments in the Dalrymple case than in the Throckmorton case?

In Brennan and Lomasky's view, our differing reactions to these two cases show that the question "what if everyone did that?" is "useful, if only as a heuristic device" for picking out when the performance of an action amounts to making an exception of oneself, or doing something unfair (2000: 77). Moreover, according to these authors, it is an imperfect heuristic device at that, since sometimes it points us in the wrong direction to an action that is perfectly permissible (such as, for example, Dalrymple's pursuit of dentistry). What the device does well, they claim, is to draw our attention to the fact that "free-riding, when generalized, is shown to be no more viable than the village in which everyone made a living by taking in others' washing" (2000: 77).

If Brennan and Lomasky are right about this, of course, the arguments I made above are very much on the wrong track. According to the view defended here, a universalization or generalization test—where you ask what would happen if everyone did something—*determines* whether an action is unfair and, when it is, purports to *explain* why it is. But if universalization tests are only imperfect heuristic devices, then, as Brennan and Lomasky write, "[strictly] speaking, what makes an ungeneralizable action wrong is not that it fails the generalization test. Rather, it fails the generalization test because of underlying unfairness, and it is the unfairness that accounts for the action's wrongness" (2000: 77). In other words, if Brennan and Lomasky are correct, I have got things exactly backwards.

Damning as the challenge may initially seem, I can resist it by amending my "what if everyone did that?" test to concern the effects of everyone's doing what you are considering doing *in the circumstances in which you are considering doing it*. In order to see why such a fix is helpful, let us begin by distinguishing between two versions of the Dalrymple case. In the first version, Dalrymple lives in a community in which there are many adults who take pride in the farming they do and many youths who look forward to a lifetime of farming themselves. In the second version, Dalrymple lives in a community in which most people who farm are unhappy about it and those people's children cannot wait to get off the farm. When the details of the original case are filled in differently like this, my intuitions diverge. In particular, it seems to me that, while in the first version of the case Dalrymple would not act unfairly by becoming a dentist, it would be unfair of her to do so in the second case. In the second case, it would make sense for the other members of her community to ask her: "Why do *you* get to be the one who leaves the farm? We do not want to be here either." In the first case, however, we can expect no such objection. For many of her fellows in the first variant, their dream is to farm.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

What this tells us is that the circumstances in which Dalrymple would pursue her dream of becoming a dentist matter to whether her pursuit of this dream would be unfair in the sense of making an exception of herself. If the circumstances in which one's action would be performed can be relevant to whether one's action would be unfair, however, then it seems natural to amend my "what if everyone did that?" test to be sensitive to the circumstances in which the action being tested for unfairness would be performed. As I alluded to above, this can be done as follows: One would make an exception of oneself by performing an action in a particular set of circumstances if, were everyone to act that way in those circumstances, then disaster would strike or an important public good would go away.

The amended version of the "what if everyone did that test?" gets the right result in each variant of the Dalrymple case. Take first, this time, the variant in which most people in her community really do not want to farm. If everyone in her community were to pursue their dream (whatever that may be), then famine would strike. That is because very few people in the world in which everyone does what Dalrymple does in the circumstances in which she does it are drawn to farming. But, then, according to the amended "what if everyone did that?" test, Dalrymple would make an exception of herself by leaving the farm for dentistry in a situation in which most farmers in her community grow fruits and vegetables begrudgingly. That is the right result. Consider now the variant of the Dalrymple case in which many people who farm enjoy doing so and many others want to follow in their footsteps by growing fruits and vegetables. If everyone in her community were to pursue their dream (whether by becoming a dentist, or a playwright, or something else altogether), then there would be no famine.

The following sort of response, however, is likely to be made on behalf of Brennan and Lomasky. Particular actions fall under multiple descriptions. To use Davidson's famous example: "I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home. Here I need not have done four things, but only one, of which four descriptions have been given" (1963: 686). When Dalrymple leaves the farm for dentistry, too, her action can be given multiple different descriptions—one of which, no doubt, is the one I have given it, namely that of "pursuing her dreams." But there are other possible descriptions. And the worry is that my amended version of the "what if everyone did that?" test will not deliver the right result about the first variant of the Dalrymple case—the one where many of her fellows want to remain or become farmers—if her action is described differently.

To see the worry, suppose that, instead of describing Dalrymple's action as "pursuing her dreams," we describe it as "becoming a dentist." Now, ask yourself: What would happen in a world that is just like the one where Dalrymple's community is full of people happy to be farming except that everyone becomes a dentist? In that world, even though many were happy to stick to farming, no one would grow fruits and vegetables anymore and the community would experience famine. This means that if we describe Dalrymple's action as "becoming a dentist," even the amended version of the "what if everyone did that?" test would deliver the wrong result about the first variant of the Dalrymple case. It would say that Dalrymple makes an exception of herself by becoming a

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

dentist, even though this is implausible when there are enough people in her community who like farming. What principled reason can we give to run Dalrymple's action through the "what if everyone did that?" test under the description of "pursuing her dreams?" The concern is that there is no such reason forthcoming.

I do, however, think that there is a principled reason to run Dalrymple's action through the "what if everyone did that?" test under the description of "pursuing her dreams"—one, moreover, that is internal to the intuitive motivation behind that very test.[5] As discussed earlier, it makes sense to ask "what if everyone did that?" when trying to figure out whether you would make an exception of yourself by performing some action. This is because, if not everyone else can do what you are considering doing without something terrible happening, then you would be privileging your own wants and desires over those of everyone else by performing that action. This kind of partiality to yourself and your own interests is a mark of unfairness. But, if this is why the "what if everyone did that?" test is a good test of unfairness, then we should only test actions under a description that would be appealing or choice worthy to most people. Otherwise, the test will not be sensitive to whether you would privilege your own interests over *those of everyone else*.

Here is why this helps my amended version of the "what if everyone did that?" test deliver the correct result in the variant of the Dalrymple case where most of her community enjoys farming. If the argument in the preceding paragraph is correct, then we should only run Dalrymple's action through the "what if everyone did that?" test under descriptions that would be choice worthy to many other members of her community. But, since, by hypothesis, most of her community prefers to farm in the first variant of the case, then we should not run Dalrymple's action through the test under the description of "becoming a dentist." Most people in her community want to keep farming, so we should not be worried about Dalrymple favoring her desires over her fellows by becoming a dentist. They would not want to become dentists anyway. This is why, I think, Brennan and Lomasky are correct that Dalrymple can successfully respond to someone who asks her "what if everyone did that?" by saying "but *not everyone will* give up farming!" And this is also why I think that Brennan and Lomasky are right that Throckmorton misses the mark when he says "but not everyone will cut across the lawn." His fellows do not take more heed of the signs than he does because shortcuts are not appealing to them. They do not cut across the lawn because they care about not being partial to themselves and their own interests over those of others in the community.

6. Other Accounts of Exception-Making

Before concluding, I would like to draw attention to one interesting feature of my account of exception-making which I have yet to discuss. My view does not take an

[5] To be clear, my claim is *not* that Dalrymple's action cannot be described as "becoming a dentist." My claim is rather that there are certain descriptions under which it makes sense to run Dalrymple's action through the "what if everyone did that?" test and certain descriptions under which it does not. And whether it makes sense for Dalrymple's action to be run through the test under a particular description can be traced to the intuitive motivation behind using that test as a test of fairness.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

agent's state of mind to be relevant to whether she would do something unfair. All that matters is whether everyone's doing what she did (regardless of for what reason she or the others did it) would result in disaster given the circumstances she finds herself in. This contrasts sharply with Wieland (2024), who has recently discussed similar cases to Dalrymple's and has offered a related analysis to my own. Roughly speaking, for Wieland, one unfairly free-rides when and because one refuses to do their part regardless of what others prefer to do (2024: 83). So, for example, I unfairly free ride by littering because I am indifferent to my neighbors in that "[I] fail to care enough about what [they] prefer"—namely, to not have to hold onto their trash until they get home (2024: 76). Now how does his view treat the two versions of Dalrymple's case? In the first version of the case in which the people other than Dalrymple do not want to farm, Wieland's view predicts correctly that it would be unfair of Dalrymple to pursue dentistry so long as her pursuit of dentistry would be done from the following maxim (M1): in order to have a fulfilling job, I will pursue dentistry regardless of what my fellow community members prefer to do for a living. Acting on such a maxim would manifest Dalrymple's indifference to her fellows, many of whom want to pursue their non-farming dreams, too.

In addition, Wieland's view correctly predicts that Dalrymple would not do anything unfair by pursuing dentistry in the version of the case in which the people other than Dalrymple *do* want to farm. At least she would not on the assumption that she pursues dentistry on the basis of the following maxim (M2): in order to have a fulfilling job, I will pursue dentistry, but only when enough of my fellows prefer to farm. If, instead, she pursued dentistry on the basis of M1 and just got lucky that her fellows wanted to farm, she would still be displaying the kind of indifference sufficient to count as making an exception. It is in *this* sense that Wieland takes an agent's mindset to be relevant to whether her action is fair or unfair. On my view, however, Dalrymple would do something perfectly fair and defeasibly permissible if she pursued dentistry on the basis of M1 so long as the others do in fact want to farm. Unlike Wieland, I find it more plausible to say that Dalrymple's pursuit of dentistry on the basis of M1 suggests a kind of character flaw or vice rather than that her action is unfair in the way that tends to make it impermissible. This difference between my account and Wieland's reflects the fact that our views of exception-making are defended in the context of normative ethical theories which part ways on the question of whether an agent's reason for acting is determinative of her action's permissibility. While contractualism, for example, denies this, the Kantian view with which Wieland is working draws a tight connection between right action and good *willing*.[6]

My account of exception-making is what Wieland (2024: 63) would call an "other-based" account, because it takes the preferences of Dalrymple's fellows rather than Dalrymple's psychology to be decisive in determining whether she would do something unfair by pursuing dentistry. Along this dimension, my account is similar to Trifan's account of exception-making, according to which I unfairly free ride on the efforts of others if they have the "free-rider's preference"—

[6] See Scanlon (2008) for a defense of the view that the permissibility of an agent's action is not to be determined (even in part) by the reason for which it is performed.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

that is, if they are willing to pay the cost of cooperation but prefer to do otherwise (2020: 167). Although I suspect that my account and Trifan's come apart in a range of cases (such as cases where other people may want to do something other than cooperate but do not have the ability to do so), our views are largely compatible. The main difference lies in our overall aim. While Trifan's central aim is to identify the wrong of free-riding and trace its contours, mine is to identify the proper role of exception-making in a general account of what makes an action permissible.

7. Wrapping Up

On my view, the question "what if everyone did that?" does have an important role to play, even if it does not play quite as *fundamental* of a role in moral theorizing as rule consequentialists and orthodox contractualists have believed it to play. In particular, I have argued that the question "what if everyone did that?" gives expression to one, and only, one kind of moral objection in our repertoire—that, by doing the thing in question, one would make an exception of oneself. Once we acknowledge this fact (which commonsense moral thought was already privy to in its use of the question), we have the resources to explain the wrong of free-riding without courting the ideal world problem. And this is true whether we are act consequentialists, act contractualists, or defenders of some other non-collectivist moral theory (such as a maxim-based one).[7]

AARON SALOMON
CHAPMAN UNIVERSITY, SMITH INSTITUTE FOR POLITICAL ECONOMY AND PHILOSOPHY
aasalomon@chapman.edu

References

Bourguignon, Lea "On the Possibility of Act Contractualism." *Australasian Journal of Philosophy*, forthcoming.
Brennan, Geoffrey and Loren E. Lomasky (2000) "Is There a Duty to Vote?" *Social Philosophy and Policy*, 17, 62–86.
Cullity, Garrett (2008) "Public Goods and Fairness." *Australasian Journal of Philosophy*, 86, 1–21.
Davidson, Donald (1963) "Actions, Reasons, and Causes." *The Journal of Philosophy*, 60, 685–700.
Hooker, Brad (2000) *Ideal Code, Real World*. New York: Oxford University Press.
Hume, David (2011) *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton. New York: Oxford University Press.
Kagan, Shelly (1989) *The Limits of Morality*. New York: Oxford University Press.
Klosko, George (1992) *The Principle of Fairness and Political Obligation*. Lanham: Rowman and Littlefield.

[7] For generous feedback, thanks to Tony Hernandez, Taylor Koles, Gary O'Brien, Japa Pallikkathayil, Ashley Purdy, Kyra Salomon, Aaron Segal, and two anonymous referees for the *Journal of the American Philosophical Association*. Thanks especially to Jussi Suikkanen for his encouragement of the overall project as well as his sensible advice about how to execute its details.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press

Munoz, Daniel (2021) "The Rejection of Consequentializing." *The Journal of Philosophy*, 118, 79–96.
Murphy, Liam (2021) "Nonlegislative Justification." In Jeff McMahan, Tim Campbell, James Goodrich, and Ketan Ramakrishnan (ed.) *Principles and Persons: The Legacy of Derek Parfit* (New York: Oxford University Press), pp. 247–276.
Nefsky, Julia (2019) "Collective Harm and the Inefficacy Problem." *Philosophy Compass*, 14, 1–17.
Parfit, Derek (2011) *On What Matters, Volume One*. New York: Oxford University Press.
Podgorski, Abelard (2018) "Wouldn't it be Nice? Moral Rules and Distant Worlds." *Nous*, 52, 279–294.
Portmore, Douglas (2009) "Consequentializing." *Philosophy Compass*, 4, 329–347.
Portmore, Douglas (2011) *Commonsense Consequentialism*. New York: Oxford University Press.
Rawls, John (1999) *A Theory of Justice*: Revised Edition. Cambridge: Harvard University Press.
Rumbold, Benedict (2024) "The Ideal Worlds Objection?" *Philosophy Compass*, 19, 4–11.
Salomon, Aaron (2024) "Maxim and Principle Contractualism." *Journal of Ethics & Social Philosophy*, 26, 1–30.
Scanlon, T.M. (1998) *What We Owe to Each Other*. Cambridge: Harvard University Press.
Scanlon, T.M. (2008) *Moral Dimensions*. Cambridge: Harvard University Press.
Sheinman, Hanoch (2011) "Act and Principle Contractualism." *Utilitas*, 23, 288–315.
Suikkanen, Jussi (2020) *Contractualism*. New York: Cambridge University Press.
Suikkanen, Jussi (2024) Act- and Rule-Consequentialism: Two Syntheses." *Moral Philosophy and Politics*, 12, 107–126.
Trifan, Isabella (2020) "What Makes Free Riding Wrongful? The Shared Preference View of Fair Play." *Journal of Political Philosophy*, 28, 158–180.
Wieland, Jan Willem (2024) "Kantian Free Riding." *Journal of Ethics & Social Philosophy*, 29, 62–85.

https://doi.org/10.1017/apa.2025.13 Published online by Cambridge University Press