
RETHINKING THE PRINCIPLE OF FAIR PLAY

BY

JUSTIN TOSI

Abstract: The principle of fair play is widely thought to require simply that costs and benefits be distributed fairly. This gloss on the principle, while not entirely inaccurate, has invited a host of popular objections based on misunderstandings about fair play. Central to many of these objections is a failure to treat the principle of fair play as a transactional principle – one that allocates special obligations and rights among persons as a result of their interactions. I offer an interpretation of the principle of fair play that emphasizes its similarities to another transactional principle: consent. This interpretation reveals that playing fair requires one to reciprocate specifically by following the rules of the cooperative scheme from which one benefits, just as consent requires one to act according to the terms of an agreement. I then draw on the comparison with consent to reply to some popular and persistent objections to the principle.

Suppose that public transportation in your town is run on an open system. Passengers are to buy a ticket at an automated kiosk before boarding the train, though no transit employee or device blocks those who do not have a ticket from boarding. Knowing a good thing when you see it, you skip the ticket-buying part of this system and board the train anyway. Have you done anything wrong? And if so, why was it wrong?¹

Some philosophers have argued that refusing to submit to the relevant rules in cases like this violates the principle of fair play.² John Rawls offers the following formulation of that principle:

... a person is required to do his part as defined by the rules of an institution when two conditions are met: first, the institution is just (or fair), ... and second, one has voluntarily accepted the benefits of the arrangement or taken advantage of the opportunities it offers to further one's interests. The main idea is that when a number of persons engage in a

mutually advantageous cooperative venture according to rules, and thus restrict their liberty in ways necessary to yield advantages for all, those who have submitted to these restrictions have a right to a similar acquiescence on the part of those who have benefited from their submission.³

The principle of fair play promises to explain the wrongness of some forms of free-riding. Return to the public transportation case. According to the principle of fair play, you have a special obligation to the other riders to submit to the same rules as they have.⁴ By withholding payment, you enjoy the benefits of others' rule-following behavior while refusing to follow the same rules yourself. You free-ride on their efforts, and so violate a fair play obligation that you owe to them.

The principle of fair play is a mainstay in many debates in moral and political philosophy. It is perhaps the most promising basis for political obligation.⁵ Some have also offered fair play accounts of state punishment,⁶ civil disobedience,⁷ global distributive justice,⁸ and stakeholder theory.⁹ Nevertheless, the principle is still widely misunderstood, and philosophers often reject it on the basis of misreadings and bad arguments.¹⁰ My aim here is to correct some of these general misunderstandings, in the hope that the principle will be applied and discussed more accurately in particular debates.

I suspect that one reason that the principle is not more widely accepted is that many (including proponents of the principle) are confused about the role of a more general form of fairness in considerations of fair play. This confusion leads to a distorted picture of what fair play obligations require, and the conditions under which fair play obligations arise. This article aims to correct that picture. I will argue that the principle of fair play is actually far more like the principle of consent than many realize.¹¹ The key point is that the principle of fair play does not require beneficiaries simply to bring about *some* fair redistribution of benefits and burdens, but rather to provide a fair return specifically by following the rules of a cooperative scheme. Thus, the rules of the cooperative scheme to which one owes a fair play obligation perform a function similar to the terms of a contract to which one has consented. I will also show that drawing parallels between fair play and consent helps to highlight features of fair play that suggest convincing replies to some common objections to the principle.

After offering a close reading of the principle of fair play and identifying its key elements, I will fill out the comparison with consent in the second section. I then consider objections that the principle of fair play collapses into either consent or gratitude in the third section. In Section 4 I respond to the argument that a beneficiary of a cooperative scheme can play fair without following the rules. Finally, I address the issue of causally ineffective rules in the fifth section.

1. *The elements of fair play*

Let us begin by identifying the conditions that give rise to fair play obligations. I will use Rawls's voluntarist formulation of the principle for these purposes, for reasons that will become clear shortly. Modifying only slightly the relevant text from *A Theory of Justice*, we will say that:

- (1) a person is required to do his part as defined by the rules of an institution (**rules condition**)
- (2) when the institution is reasonably just and fair (**justice condition**),
- (3) and one has voluntarily accepted (**acceptance condition**)
- (4) the benefits of the arrangement (**benefit condition**).

One further condition is made clearer in Hart's formulation and Rawls's original discussion of the principle:

- (5) 'cooperation requires a certain sacrifice from each person, or at least involves a certain restriction of his liberty'¹² (**costs condition**).

The conjunction of these five conditions is necessary and sufficient for the generation of fair play obligations. I have omitted some elements from the various formulations that I take to be superfluous.¹³ The elements listed above are the core of the principle of fair play.

I will now make some preliminary remarks to elaborate on each condition. Many will also be fleshed out more as I consider objections further down. I am less interested here in defending these conditions than I am in laying out some of the main details of fair play. One thing that will become clear is that there is no monolithic fair play view about what some of the thick moral terms (e.g. 'just and fair', 'voluntarily', 'benefits') in the principle of fair play mean. There are, instead, a family of possible conceptions of fair play. Fair play theorists need not come to an agreement on what the precise criteria for fairness are, or what counts as a benefit, any more than consent theorists must agree on what counts as duress, or compromised agency, or unconscionability in the terms of a contract.

Take first the rules condition. There are two necessary conditions for the existence of a rule. First, rules must be regarded as requirements that those bound by them *ought* to submit to, and not merely as guides for predicting behavior. Thus, failure to comply with a rule must be thought of as grounds for blame or criticism, rather than a curious aberration in a pattern. Second, rules must be recognized as existing – otherwise they could not fulfill their action-guiding function. As Hart famously argued, this recognition can come in either of two forms: in the rule being practiced (as in systems of primary rules), or in its validity according to the scheme's rule of recognition (as in modern legal systems).¹⁴

The justice condition consists of two parts. First, the scheme must be just in the sense that it respects our natural duties of justice (more on this in the next section). And second, the benefits and burdens of the scheme must be fairly distributed among its members.¹⁵ There is nothing about the principle of fair play that commits a fair play theorist to any specific view of the substantive requirements of justice or fairness. But we can usefully distinguish between two broad views about how justice limits the possibility of fair play obligations. Some prefer a strong justice condition, so that the introduction of any injustice in the ends of the scheme, or any imbalance in the distribution of benefits and burdens, undermines the possibility of genuine fair play obligations.¹⁶ Others hold, more commonly, that the distribution of benefits and burden must be reasonably fair, and the rules of the scheme reasonably just.¹⁷ A scheme need not be optimal in order to provide us with a set of benefits that is well worth the cost. Thus, there is some threshold of injustice or unfairness past which fair play obligations are extinguished, but there is some space for suboptimal rules between that threshold and perfection.

I have included Rawls's acceptance condition rather than Hart's mere receipt condition mostly in order to present the most accessible versions of the principle. I think that mere receipt is sometimes sufficient as a commitment mechanism for fair play obligations, but a defense of that view requires raising other contentious issues in moral philosophy. What does it mean to accept a benefit? The clearest instances involve persons taking benefits willingly, and with the knowledge that they are the products of a cooperative scheme conducted according to rules – as in the public transportation case. Whether a person accepts a benefit is less clear in the case of non-excludable goods, as there is generally no observable sign of the recipient's attitude regarding receipt.¹⁸ But the sense of acceptance that interests us has primarily to do with what the person wills. In this sense, acceptance is a psychological state, not an outwardly observable performance. So what matters for purposes of fair play is the recipient's attitude itself, not her observable behavior. If she has a pro-attitude toward receipt, knows of the scheme's rules, and the benefits are worth the cost of submitting to the rules (whatever the correct account of costs and benefits turns out to be), then she incurs a fair play obligation.¹⁹

Like the justice condition, the benefits condition raises controversial substantive issues in moral philosophy. If one is a desire-satisfaction theorist about well-being, then whether something constitutes a benefit depends on whether the recipient (or some suitably idealized version of the recipient) desires it. Whether it is worth the cost depends on the recipients' valuation of the benefit relative to the rule-imposed cost. Objective list theorists, on the other hand, may think that (at least for some goods) there is a fact of the matter about whether something is a benefit and whether it is worth the cost.²⁰ Again, one can be a fair play theorist and hold any of these views.

There is no need to insist that just one of them is bound up with the principle of fair play.

The same comments about well-being apply straightforwardly to the costs condition. It is essential that participants in the scheme incur costs in supporting it, because otherwise there is no debt of fairness for beneficiaries to repay. A beneficiary may show a lack of gratitude by not reciprocating in some way, but she does not treat those who benefit her unfairly.

2. *Fair play, fairness, and consent*

Ronald Dworkin writes:

... there is no general moral principle that requires me to contribute to the cost of producing what benefits me: I may be selfish when I pass a street musician by without tossing him a bill, but I violate no obligation even if I have enjoyed his music – even if I have paused to hear more of it.²¹

Dworkin is right about one thing. A general principle requiring reciprocity in all cases like the one he describes would have absurd consequences. Positive externalities of others' actions are all around us, yet we do nothing wrong by enjoying these free benefits of social life without contributing to their production.²²

Fortunately, denying the existence of such a principle does not interfere with the foundations of the principle of fair play. Although some philosophers instead refer to it as the principle of fairness, fair play is a better term for the central moral concern captured by the principle. The name has proved to be a source of confusion, as it suggests a more prominent role for broader considerations of fairness (or, relatedly, distributive justice) than the principle actually allows. Fairness is a more general notion than is fair play, and violations of fair play are not the only kind of unfairness. At the most general level, fairness is a kind of impartiality between persons. Political philosophers have lately been most interested in fairness concerning the distribution of goods between persons. Thus, it is understandable that a 'principle of fairness' would seem to many to amount to the idea that any imbalance in the distribution of benefits and burdens must somehow be corrected.²³ While the distribution of benefits and burdens clearly plays a role in the principle of fair play, that role is far more limited than some have suggested.

To play fair with others is to treat oneself as equally subject to the same set of rules as them. Playing fair is, then, a more localized concern than fairness more generally, as it occurs within the bounds of a cooperative scheme conducted according to rules. Playing fair is not about the distribution of benefits and burdens between people in general, or even with respect to all their interactions. Thus, in a high school basketball game, the players may play

fair even if one school is from an unfairly wealthy district and the other is from an unfairly poor district. But while the background conditions that are not governed by the rules of the game need not be fair, fair play is not possible if the rules of basketball are unfair. An uncomfortable team bus is irrelevant for concerns of fair play, but a smaller rim for one of the teams is. So, playing fair does require a concern for distributive justice with respect to the benefits and burdens of the cooperative scheme, but the scope of the concern is no wider than that.

Since the imperative to play fair is a narrower requirement than fairness broadly understood, perhaps we would do well to find another way of thinking about how the principle works. I think that we can best understand the principle of fair play by comparing it to the principle of consent, and asking what work is done by each part of the principles. It should be unsurprising that they are similarly structured, as both are transactional principles: they explain the allocation of special obligations and rights among persons by appeal to facts about some interaction among those persons. John Simmons classifies obligations arising from consent as ‘obligations of commitment’ – i.e. ‘obligations deliberately undertaken’ – and those generated by the principle of fair play as ‘obligations of reciprocation’ – i.e. obligations ‘generated by the receipt or acceptance of benefits.’ As I understand them, fair play obligations fit under the first heading as well as they do the second.

Obligations are incurred through consent by the giving of a sign of one’s will to agree to a set of specific terms, according to which rights are then allocated to the parties to the agreement.²⁴ The two main functions to notice here are specification and commitment. The set of terms, or contract, does the work of specifying the rights to be conferred to the parties. And the parties are committed to those terms by the expression of their will. That expression may, of course, be express or tacit, explicit or implicit. What matters is that it is a reliable signal of the intention of the consenting party.

We can locate the same functions in parts of the principle of fair play. The work of specifying the terms of the relationship is done by the rules of the cooperative scheme from which one accepts benefits. Since rules can take the same forms as can the provisions of a contract specifying the rights to be allocated to the parties, fair play can generate the same allocation of rights as any contract. As we have already seen, acceptance (or, according to some, mere receipt) of the scheme’s benefits is the commitment mechanism of fair play. And again, though acceptance may be read into persons’ behavior (e.g. going out of one’s way to get benefits), there need not be any reliable social signal of acceptance as there must be a reliable sign of consent.

These are the most important structural similarities, but there is more to the comparison between fair play and consent. Each principle offers various kinds of protection against people becoming committed to bad arrangements. Consent is valid only if the parties are sufficiently informed about what they are agreeing to, and fair play requires promulgation of the rules

to which those who accept benefits must submit. Consent to unconscionable contracts is said by most consent theorists to be nonbinding. The analogue for fair play is that one cannot be required to submit to an unfair set of rules.

But what about unjust rules? Simmons argues that Rawls's stipulation that a cooperative scheme must be just in order to generate fair play obligations is unmotivated.

The intuitive force of the principle of fair play seems to be preserved even for criminal conspiracies, for example. The special rights and obligations which arise under the principle are thought to do so because of the special relationships which exist between the cooperating participants; a fair share of the burdens is thought to be owed by a benefiting participant simply because others have sacrificed to allow him to benefit within a cooperative scheme. No reference is made here to the morally acceptable status of the scheme. Simple intuitions about fair play, then, do not seem to provide a reason for disqualifying unjust cooperative schemes.²⁵

It is interesting to reread this passage while imagining that Simmons is criticizing consent rather than fair play.²⁶ After all, nothing about the idea of making an agreement with another person suggests that the agreement must concern morally acceptable actions. Suppose the hitman that Jones hires refuses to carry out his part of the contract after being paid. Does Jones have any moral complaint for non-performance arising from the hitman's consent to the contract?²⁷ This question could be answered in one of two ways.

On the one hand, some moral philosophers would conclude that Jones has no moral complaint for non-performance despite having obtained the hitman's consent to the deal. The deal confers no moral rights to either party, on this view, as its terms do not have the sanction of morality. Jones can have no claim right that the hitman murder the target, because the hitman can have no liberty right to do so. On the other hand, one could hold that the agreement does generate a *pro tanto* obligation for the hitman. However, that obligation has to compete with all of the hitman's conflicting moral requirements, and it is clearly overridden by his much weightier *pro tanto* duty not to murder.

The important thing to see is that consent and fair play are in the same boat with respect to this issue. All of the same content restrictions present in cases of consent also apply to the possibility of the rules of a cooperative scheme being transformed into moral requirements by the principle of fair play. If I benefit from the rule-based heist that I pull off with my partners, whether they have a moral basis for complaint (on grounds of fair play or otherwise) if I make off with their share of the loot depends on how transactional moral principles like fair play and consent work in general. If I owe fair submission to the rules to my fellow heisters as a matter of fair play, the hitman owes fulfillment of the contract to Jones because he offered his consent. So, if this is not an insurmountable difficulty for the principle of consent, then it is not an insurmountable difficulty for fair play, either.

What we have in the principle of fair play, then, is the idea that one can bind oneself to submit to a set of rules by accepting the benefits that are attached to them as a matter of fairness. To play fair is to follow the rules of the cooperative scheme to which one owes a debt of fairness. Submission to the rules is the only return that qualifies as fair, just as doing one's part as required by the terms of an agreement is the only way of keeping to an agreement to which one offered one's consent. I will stop short of claiming that the principle of fair play is a variant of consent, but the two principles are clearly related. They address similar moral concerns through slightly different means, as is fitting for the different contexts in which they apply. And they work against similar background moral constraints about the creation and allocation of rights. As I consider common objections to the principle of fair play, I will revisit these similarities.

3. *Collapse objections*

Critics of the principle of fair play have also pointed to similarities between it and other moral principles. In fact, some argue that the similarities are so striking because so-called fair play obligations actually arise from the application of those other principles, and not some separate principle of fair play. Some claim that circumstances leading to fair play obligations are morally indistinguishable from those leading to gratitude obligations. Others argue that the principle of fair play is not only similar to consent in structure, but that it collapses into consent. The features of fair play brought into relief by the comparison above help us to see that both sets of critics are mistaken.

Let us first consider the charge that fair play is indistinguishable from gratitude. Daniel McDermott offers two main reasons to doubt that fair play covers a distinct area of morality. First, McDermott argues that the fact that goods come from a cooperative scheme, rather than an individual, is not morally significant. McDermott considers a popular example, first given by Simmons, of a well that was dug and maintained by a cooperative scheme. A group decides at a community meeting to dig a well, and a number of them pledge to do their part in supporting the project. According to McDermott and Simmons, the principle of fair play holds that if a person takes water from the well, but did not join at the meeting or bear any of the costs of supporting the project, she has incurred a fair play obligation to submit to the costs.²⁸ But, McDermott adds, suppose that it were not a cooperative scheme that dug and maintained the well, but a single private owner. Then it still seems that latecomers who take from the well owe the owner some compensation. But it cannot be for reasons of fair play, since there is no cooperative scheme to which such persons might owe submission. There are, however, broader considerations of gratitude that are present in

both the individual owner and cooperative scheme variants of the case, and these seem to explain the obligation to reciprocate (or compensate). So, fair play does not cover an area of morality separate from that already covered by gratitude.²⁹

McDermott is right that the relevant moral phenomenon does not change between the two variants of the well case. But that poses no problem for the principle of fair play, because latecomers violate no fair play obligations in Simmons's original case, either. Taking water from the well in the conditions Simmons describes cannot be a violation of fair play, because there are no rules in place providing that the benefits are open to new members who submit to some set of rules.³⁰ One might insist that latecomers are still free-riding, even if they violate no fair play obligations. I have no objection to this characterization, because I do not claim that every instance of free-riding is a violation of fair play.³¹ Latecomers may owe compensation for violating the scheme or individual owner's property rights, but they do not owe anything as a matter of fair play, for they had no means of joining the scheme under the rules. On the other hand, if there were rules providing for the required submission of those who accept benefits in both variants of the case, then anyone who took water from the well would owe fair play obligations, regardless of how many people were already supporting the provision of benefits.

McDermott's second (and related) objection is that receiving benefits from a cooperative scheme seems no different from receiving a gift. Gifts carry with them no obligation of reciprocity to the giver. Otherwise they would be an imposition rather than a gift. It would be in keeping with the norms of friendship (or perhaps an appropriate showing of gratitude) to make a proportionate gesture in return, but the gift giver has no right to such a return. So, McDermott concludes, it seems ad hoc to hold that benefits from a cooperative scheme are any different.

The rules condition again allows for a ready response to this objection. The difference between the gift context and contexts in which fair play obligations are generated is in the rules in play. The practice of giving and receiving gifts is not governed by rules of reciprocity that could be transformed into moral obligations by fair play. So there *is* a morally relevant difference between fair play cases and the case of gift-giving: one concerns the morality of rule-following, and the other does not. It could be objected that there are indeed rules surrounding the practice of giving and receiving gifts. But if this were so, it seems obvious that the rules would be very different than those in cooperative schemes that require repayment of a debt. If critics want to maintain that gifts are actually impositions if receipt of one requires repayment, then they cannot sensibly hold that there is a rule requiring repayment. And if there is no rule requiring repayment for gifts, then the principle of fair play does not generate an obligation of repayment for receiving a gift. So much, then, for the inseparability of fair play and gratitude.

Some argue that examples commonly thought of as paradigmatic cases in which fair play obligations are generated are better understood as examples of consent. The most extreme version of this argument concludes that the principle of fair play is not an independent moral principle at all, but rather that it collapses into consent upon investigation. Consider again the public transportation free-rider case discussed at the outset. One might agree that there is an obligation to pay the fare in cases like this, yet deny that fair play is the basis of that obligation.

Jiafeng Zhu offers a recent statement of this collapse argument.³² His alternative explanation for the passengers' obligation to pay the fare in the public transportation case is that they offer their tacit consent to pay by riding without paying. The free-rider tacitly consents to pay the transit fare by boarding, Zhu says, just as one tacitly consents to pay one's bill in a restaurant by placing an order.³³ If the system were run differently, with ticket sellers blocking entrance to the vehicle, riders would expressly consent by paying. But since the system is open (i.e. little to no effort is made to exclude people from enjoyment of the good before payment), riders tacitly consent to pay by boarding the train. In Zhu's view, using the open system instead of a ticket seller 'does not change the nature of the public transport scheme; it only makes the system more open, and changes the way people pay their fares.'³⁴ Thus, he concludes, the best explanation of the case is that the free-rider violates a tacit agreement by failing to pay.

Zhu's account of the duty to pay the fare rests on a mistake about the nature of tacit consent. One offers tacit consent by remaining silent or inactive in cases where doing so is a sign in the given context that one consents.³⁵ As Zhu has it, the behavior of a free-riding passenger displays a reliable conventional sign to others that she agrees to pay the fare (and follow any other relevant rules of the public transportation system). But this is implausible. Short of an oddly explicit spoken declaration that one intends to ride for free, there could be no more reliable sign that one does not consent to pay than walking past the ticket purchasing stations and boarding the train without paying. If the behavior of a free-riding passenger qualifies as a sign of consent, then it would seem nearly impossible to avoid signaling consent.

It would be more plausible to think that the free-rider incurs an obligation to pay the fare regardless of whether she offers a sign of consent, and the principle of fair play provides an appealing explanation for this result. By accepting the benefits of the public transportation system, the free-rider becomes a member of the cooperative scheme of riders. However, she refuses to submit to the rules of the scheme by skipping out on the fare, and so claims a greater liberty for herself than that enjoyed by the other members – whose payment makes possible the benefit in which she takes part. Submitting to the rules by paying to ride is the only way she can play fair with her fellow passengers. Unless she pays, she violates a fair play obligation that

she owes to them. It seems, then, that Zhu has mistaken what Simmons calls a 'consent-implying act' with a genuine sign of consent. Boarding without paying does not signal that the free-rider intends to consent, but it does morally obligate her to pay for other reasons.³⁶

Now, there are cases in which one *does* conventionally signal one's consent by accepting benefits. (Think, for instance, of 'enter at your own risk' notices.) In these cases, consent is a fine explanation of the obligations that arise through the conventionally appropriate behavior. But the existence of such cases will not save Zhu's collapse argument, for two reasons. First, acceptance of benefits is not always a sign of consent, even if it is sometimes. Yet it seems clear that one may incur an obligation to follow the rules in such cases, anyway. For Zhu's argument to succeed, he would have to show that acceptance *always* signifies consent in the relevant cases, or that the obligations incurred in these cases are based on some moral principle other than fair play. But the public transportation case is a clear example in which acceptance of benefits is not a sign of consent, and fair play is a strong candidate explanation for the relevant obligation.³⁷

Second, even if consent can show that a person is under an obligation in some case, that obligation may be overdetermined. In other words, one could be bound to fulfill some obligation both because one consented to do so, and as a matter of fair play.³⁸ Take a case involving another kind of special obligation. If my mother is sick, I have an associative obligation to her to provide for her care. But suppose that, for some reason, I once promised a friend that I would take care of my mother should the need arise. Then I owe performance of the same actions to him for different moral reasons. If one were to declare one of these grounds invalid – out of a concern for simplicity, say – one would miss morally salient features of my resulting moral requirement. The multiple grounds of obligation in play in this case illuminate the fact that I owe the provision of care to different people for different reasons. The same could be true of some private variant of the public transportation case. I may offer my consent to follow the rules of the train (pay my fare, keep my clothes on, etc.) to the train company, and also owe the same thing to my fellow passengers as a matter of fair play.

Although fair play obligations are based on moral considerations similar to those that generate obligations of gratitude and consent, we can now see clearly that fair play does not collapse into either of these other principles. Even though fair play and gratitude are both to some degree obligations of reciprocity, the role of rules in the principle of fair play distinguish it as a separate moral principle. And similarly, while consent and fair play each have elements of commitment, the commitment mechanism of fair play allows it to generate obligations in cases where the commitment mechanism of consent (a sign of one's will to agree) is not present.

4. *Why follow the rules?*

Playing fair with others requires doing one's part to fairly distribute the benefits and burdens of cooperation within a joint enterprise. The principle claims that one must follow the scheme's rules in order to do one's fair share. But some critics have claimed that one might bear a fair share of burdens in any number of ways.³⁹ One critic even suggests that one is not free-riding if one is taking a severe enough risk of being caught and subjected to punishment, since risk is a cost.⁴⁰

Zhu has also recently questioned the fair play rationale for rule-following. He argues that fair play theorists have moved too quickly from a commitment to play fair to a requirement of submission to the rules, and so leave a 'justificatory gap.'⁴¹ Though he concedes that the principle of fair play may generate an obligation for beneficiaries of a cooperative scheme to reciprocate to its participants in some way, he claims that it cannot bind beneficiaries to repay specifically by cooperating – i.e. following the scheme's rules. As Zhu points out, fair play theorists need some way of justifying the claim that beneficiaries do not retain a liberty right to choose how they repay their debt of fairness.

Fair play theorists have an obvious initial reply. They may argue that discretion in how one chooses to discharge a debt is a benefit. Part of the cost borne by the rule-following members of the scheme is forfeiture of the liberty right to choose their form of repayment. Thus, the fair share of the cost for new members of the scheme includes repayment specifically by submission to the rules. But critics have an equally obvious reply: if it is an additional cost to submit to the rules rather than reciprocate freely (i.e. in a form of one's choosing), then one could still pay one's debt of fairness to the scheme through an alternative payment of equivalent value that includes that additional cost. In other words, one could retain the liberty right to choose the form of repayment, and compensate the scheme for that extra benefit by bearing a different burden of similar value. So, this initial reply will not do. The fair play theorist needs some other rationale for the requirement that one repay specifically by following the rules.

Here again we can better understand the principle of fair play by comparing it to consent. Notice that this objection – why follow the rules? – has the same general form as asking why one ought to perform one's part of a contract as the terms specify, rather than instead compensating others with some payment of equivalent value. In other words, the objection aims to undermine fair play's specification mechanism. But the same objection, when considered against consent, is plainly absurd. Few would be moved, I imagine, by an objection to consent theory that the parties to an agreement owe *something* to one another, but that what they agree to is simply one thing among many that might satisfy their obligations. The terms of an agreement are not merely a suggestion about what one might do for the other party, just

as the rules of a cooperative scheme are not merely a suggestion. The important thing to see, again, is that consent and fair play are in the same boat. What is at stake is the ability of transactional principles to specify the content of obligations, rather than merely establish that some amount of value is owed.

Let us consider one possible argument for following the rules that makes the consent comparison more apparent. Rawls writes: 'The reason one must abstain from [free-riding] is that the existence of the benefit is the result of everyone's effort, and prior to some understanding as to how it is to be shared, if it can be shared at all, it belongs in fairness to no one.'⁴² Rawls's point is that the benefit that the scheme produces does not belong to any of its members individually, since it was no one person's effort that produced it. Rather, the benefit is held by the scheme as a collective agent, and can only be enjoyed fairly by individuals when the terms of that enjoyment are established by conventional rules. In the absence of rules the benefits may not be owned, and therefore not enjoyed, by any one person.⁴³ The rules of a cooperative scheme, then, are the terms according to which collectively held property is made accessible to individual members.

This argument may be rendered slightly less opaque by again thinking of fair play in terms similar to consent. The benefit produced by the scheme is jointly owned by default. The scheme's rules concerning the distribution of burdens and benefits are a kind of open offer, setting the terms for access to the jointly held property. They say that an individual may enjoy the scheme's property only by submitting to the rules, and so acceptance of benefits is tied to submission to the rules. If one simply takes the scheme's benefits and reciprocates freely, then it is as if one claims a right to bypass the part of the transaction in which the wishes of the other party are considered. The free reciprocator ignores the scheme's stated offer, imposes her own will on others, and unjustly takes property to which she has no right. She is better than the free-rider only in that she offers compensation for her rights violation, though that compensation is also unilaterally imposed.

5. Fair play and causal contribution to benefits

Michael Huemer presents a related objection about rule-following. Intuitively, fair play is about doing your part to produce a group benefit, part of which you enjoy. According to fair play theorists, this means following the rules. But many cooperative schemes have rules that make no causal contribution to the production of benefits at all. Huemer argues that it could not be unfair to refuse to follow some individual rules, as they may turn out to be either ineffective or counterproductive to the ends of

the cooperative scheme. He proposes the following revised conditions for the principle of fair play:

- i) The scheme is producing a significant good.
- ii) The other participants are bearing a cost that is 'causally necessary' to the production of that good.⁴⁴
- iii) You receive a fair share of the good being produced.
- iv) Your contribution would causally contribute to the production of the good.
- v) The costs you bear in participating would be reasonable and not significantly greater than those borne by others.
- vi) Your participation in the scheme would not interfere with your doing something more important.⁴⁵

These conditions allow Huemer to give what he takes to be the intuitively correct assessment about whether it is unfair to disobey causally ineffective rules. His answer, as these conditions clearly suggest, is 'no.' Huemer describes a lifeboat case to illustrate his thinking, and to provide a not-so-subtle allegory for contemporary democratic politics:

The lifeboat is taking on water. The passengers gather and discuss what to do about the problem. A majority (not including you) want Bob to devise a solution. Bob thinks for a minute, then announces the following plan:

- i) All passengers shall start bailing water out of the boat;
- ii) They shall pray to Poseidon to ask for his mercy;
- iii) They shall flagellate themselves with belts to prove their seriousness;
and
- iv) They shall each pay \$50 to Sally, who helped Bob get elected.

You know that item (i) is useful, item (ii) useless, and items (iii) and (iv) harmful to most passengers. Nonetheless, most other passengers participate in all four parts of Bob's plan. If you refuse to pray, self-flagellate, or pay Sally, do you thereby act wrongly? Do you treat the other passengers *unfairly*?⁴⁶

Again, on Huemer's view, fair play requires only that one follow the first rule. The view is attractive for obvious reasons. It seems pointless to obey ineffective rules. Requiring obedience from those who can see that following a rule will be wasteful is insulting and frustrating. And of course, one could easily imagine even more ridiculous rules than those Huemer lists.

But although Huemer's conception of fair play is within the family of possible fair play views, it is not among the more plausible conceptions. To see why, consider a similarly constrained version of the principle of consent. Suppose we held that consent to follow a rule is invalid when the rule is causally ineffective in promoting its end. As with Huemer's version of fair

play, this is a possible view, and from a certain ideal standpoint both views make perfect sense. But that ideal standpoint is not one that is useful for solving problems that require cooperation between non-ideally situated persons.

Here is Huemer's advice about disobeying unjust laws:

If a law is *unjust*, one may break it. But it is not the case that if one merely believes a law to be unjust, one may break it; it depends upon whether one's belief is correct.

There are many cases in which we cannot tell whether a law is just or unjust; justice is a difficult subject. What ought we to do then? In cases where we do not know whether the law is just, we will simply not know whether it is permissible to break that law. I can say nothing here that will cause readers to be able to know in all cases what is just, or what they ought to do. My only advice for such situations is that one do further research on the topic (perhaps in the ethical and political philosophy literature), and then exercise one's best judgment.⁴⁷

The same advice is straightforwardly implied by the causal effectiveness condition. There is no obligation to follow the rule if it does not causally promote its stated goal. But here is the trouble. Sometimes, in order to secure the cooperation of others in important joint enterprises, we have to agree to do things that we may think are pointless wastes of time and resources. But merely being able to utter words of agreement is not enough. That is because if, in uttering those words, we do not actually form valid moral commitments to follow rules that we think are ineffective, then neither do the persons with whom we must compromise in order to solve problems.

Return again to the lifeboat case. Suppose that no one will comply with anyone else's favored rule unless she has a moral assurance that others will comply with her own favored rule, and that the boat will sink without multiple people bailing water out. (Stipulate, if you like, that non-compliance with any rule will not be detected – until, I suppose, the lifeboat sinks.) Here both the constrained principle of consent and Huemer's version of fair play will be disastrous. As the ship sinks, the passengers will all wish that they could have committed themselves to follow the rules that they thought were ineffective. Of course, Huemer might reply, they could just as easily wish that they had been stuck on a lifeboat with people who all knew how to keep the thing from sinking. But then again, we might all wish the same thing about everyone in the world with whom we have to cooperate in order to solve problems. Suboptimality is a hallmark of human relationships. Huemer's causal effectiveness condition, then, has the result of leaving us helpless in the face of a common source of prisoner's dilemmas – disagreement about the effective means of solving problems. But this suggests that something has gone wrong, as one of the main points of transactional moral principles like consent and fair play is that they give us the ability to commit to agreements with others in order to assure them that we are willing to

compromise on our second or third choice for the sake of successful cooperation. Whatever else is true of morality, it is not a suicide pact. It does not leave us helpless in the face of threats that could have been easily avoided with something as simple as the ability to commit ourselves to non-ideal agreements or sets of rules. So, while Huemer's constrained principle of fair play is a conceptual possibility, I think it is best thought of as no more than that. We can be required to follow the ineffective rules of a functioning cooperative scheme as a matter of fair play. Such requirements may seem pointless at first glance, but they are not. Individual rules are only effective in producing benefits when they are followed. What matters for the principle of fair play is that the total system of a scheme's rules is causally effective.

6. *Conclusion*

Many objections to the principle of fair play rely on misunderstandings about what sort of moral principle it is. It is common to think of fair play as a requirement of general distributive justice, with the fair play that it requires apparently being the same thing as fairness in general. I have tried to show that this is a mistake, and that once we reinterpret the principle as being much like other transactional moral principles – in particular, consent – we are well-positioned to see that it is far more plausible than critics have claimed.

This article has focused almost entirely on the critics of fair play, so I will conclude by summarizing the good and bad news for fair play theorists. The good news is that many of the objections that have been leveled at the principle for the past several decades can be answered simply by either explaining the role of rules in specifying the content of fair play obligations or drawing comparisons between fair play and consent. The cost of this firmer footing is the bad news that it may be more difficult to meet the conditions for fair play obligations than some theorists have thought. In particular, fair play theorists will have to do more than point to the production of benefits by some person or institution to establish that those who receive or accept those benefits owe anything to the producer as a matter of fair play. They will have to show that the benefit-producing scheme is conducted according to rules that specify the content of obligations for new members of that scheme. This may mean that the principle cannot accomplish some of the tasks that it has been assigned. But it also means that the principle can more credibly address the problems for which it is well-suited.⁴⁸

McDonough School of Business
Georgetown University

¹ Garrett Cullity considers a similar case: Cullity, G. (1995). 'Moral Free Riding,' *Philosophy and Public Affairs* 24(1), pp. 3–34, at p. 6.

² Important defenses of the principle of fair play include Arneson, R.J. (1982). 'The Principle of Fairness and Free-Rider Problems,' *Ethics* 92(4), pp. 616–633; Cullity, 1995 op. cit.; Hart, H.L.A. (1955). 'Are There Any Natural Rights?,' *Philosophical Review* 64(2), pp. 175–191; Klosko, G. (2004). *The Principle of Fairness and Political Obligation*, rev. edn. Lanham, MD: Rowman & Littlefield; Simmons, A.J. (2001). *Justification and Legitimacy: Essays on Rights and Obligations*. New York: Cambridge University Press, pp. 1–26. Note that although Simmons is a prominent critic of the fair play account of political obligation, he does defend the principle itself.

³ Rawls, J. (1999). *A Theory of Justice*, rev. edn. Cambridge, MA: Belknap Press of Harvard University Press, 1999, p. 96. The final sentence comes very close to H. L. A. Hart's formulation, which is primarily distinguished from Rawls's in requiring mere receipt – and not voluntary acceptance – of benefits in order to generate fair play obligations. See Hart, 1955, op. cit., p. 185. Less well-known is C. D. Broad's earlier consideration of an 'argument from fairness.' See Broad, C.D. (1916). 'On the Function of False Hypotheses in Ethics,' *International Journal of Ethics* 26(3), pp. 377–397 at p. 387.

⁴ Special rights and obligations are those that arise out of special transactions or relationships between persons (e.g. contracts). They are to be contrasted with general rights and obligations, which are held against (or owed to) everyone. See Hart, 1955, op. cit., p. 183.

⁵ Dagger, R. (1997). *Civic Virtues: Rights, Citizenship, and Republican Liberalism*. New York: Oxford University Press, pp. 69–70; Klosko, 2004 op. cit.; Song, E. (2012). 'Acceptance, Fairness, and Political Obligation,' *Legal Theory* 18(2), pp. 209–229.

⁶ Dagger, R. (1993). 'Playing Fair with Punishment,' *Ethics* 103(3), pp. 473–488; Dagger, R. (2008). 'Punishment as Fair Play,' *Res Publica* 14(4), pp. 259–275; Hoskins, Z. (2011). 'Fair Play, Political Obligation, and Punishment,' *Criminal Law and Philosophy* 5(1), pp. 53–71.

⁷ Delmas, C. (2014). 'Political Resistance: A Matter of Fairness,' *Law and Philosophy* 33(4), pp. 465–488; Pasternak, A. (2017). 'Fair Play and Wrongful Benefits,' *Journal of Moral Philosophy* 14(5), pp. 515–534.

⁸ Meketa, I. (2015). 'Honor Among Thieves,' *Ethical Theory and Moral Practice* 18(2), pp. 385–402.

⁹ Phillips, R.A. (1997). 'Stakeholder Theory and a Principle of Fairness,' *Business Ethics Quarterly* 7(1), pp. 51–66; Phillips, R.A. and Reichart, J. (2000). 'The Environment as a Stakeholder? A Fairness-Based Approach,' *Journal of Business Ethics* 23(2), pp. 185–197.

¹⁰ Many still cite Robert Nozick's criticisms as the last word on the principle of fair play. I will not discuss Nozick's critique in any detail here, as many others have offered effective responses, and I am more interested in a different class of objections. Nozick's main charges are that benefits could simply be foisted on people against their will, thus apparently creating fair play obligations, and that the repayment demanded by the rules might exceed the value of the benefit provided. To the first charge, fair play theorists have either argued that benefits must be voluntarily accepted, or that they must be among a class of highly valuable goods. To the second, they note that a scheme cannot fairly demand a return greater in value than the benefit provided. Nozick, R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books, pp. 90–95. For examples of replies to Nozick, see Simmons, 2001, op. cit., pp. 11–23; Klosko, 2004, op. cit., pp. 33, 37–48.

¹¹ To avoid some initial confusion, let me note at the outset that despite the comparison to consent, I am not taking sides in the debate over whether benefits must be voluntarily accepted or may instead be passively received in order to generate fair play obligations.

¹² Rawls, J. (1999). 'Legal Obligation and the Duty of Fair Play,' in *Collected Papers*, S. R. Freeman, ed. Cambridge, MA: Harvard University Press.

¹³ For instance, both Hart and Rawls write that the principle may apply when 'a number of persons' engage in a joint enterprise. This wording is sometimes interpreted to mean that an enterprise must have multiple members before others can be obligated to submit to the rules out of considerations of fair play. Whether multiple members are necessary depends on a technical point about the nature of rules. If a single person can create and submit to a set of fair rules on his own, and in doing so provide benefits to others, then multiple members are not necessary. If, on the other hand, rules can only exist concerning the conduct of multiple people, then multiple members are necessary.

¹⁴ See Hart, H.L.A. (1994). *The Concept of Law*, 2nd edn. Oxford: Clarendon Press, pp. 9–11; Shapiro, S. (2011). *Legality*. Cambridge, MA: Harvard University Press, p. 90. These existence conditions are drawn from Hart.

¹⁵ As Simmons notes, it is not clear whether failing to meet this second requirement undermines fair play obligations for the entire scheme, or only for those who do not enjoy a fair distribution. This issue seems to me to hinge on whether unfairness to some members can be construed as a cost to others. If so, and if the cost is sufficiently high, then the unfairness of the distribution would undermine fair play obligations for all members. Simmons, 2001, op. cit., pp. 5–9.

¹⁶ For example, Richard Arneson has recently defended the view that the rules of a cooperative scheme must be Pareto-optimal in order for the participants of a scheme to have fair play obligations: Arneson, R. (2013). 'Paternalism and the Principle of Fairness,' in C. Coons and M. Weber (eds) *Paternalism: Theory and Practice*. New York: Cambridge University Press, p. 141.

¹⁷ Rawls, *A Theory of Justice*, op. cit., pp. 308–12. Rawls seems to be of two minds about the justice condition. In the passage just cited he argues that we owe our fellow citizens a 'duty of civility' (not to be confused with a different duty with the same name discussed in his later work) to comply with laws that treat us unjustly, as long as the burdens of injustice are not too great in any particular case and are evenly distributed in the long run. But elsewhere he describes a much stronger justice condition for the principle: 'the institution is just (or fair), that is, it satisfies the two principles of justice' (ibid., p. 96.). It is not clear how this statement can be squared with Rawls's claim that the principle of fair play is the basis of all special obligations. (How could the institution of promising satisfy the difference principle, for instance?)

¹⁸ A good is non-excludable if it is impossible or inefficient to exclude people from enjoying it.

¹⁹ What if she does not know of the scheme's rules? If the rules are not well-publicized, then she can hardly be thought to act unfairly in willingly accepting what appears to be a freely available benefit without submitting to any costs. But people are sometimes culpably ignorant about the rules because they willfully maintain false beliefs, or avoid sufficiently looking into the matter. When lack of knowledge is attributable to a moral failing, some hold that fair play obligations may arise anyway. See Arneson, 1982, op. cit., p. 632; Renzo, M. (2014). 'Fairness, Self-Deception and Political Obligation,' *Philosophical Studies* 169(3), pp. 467–488. For a more detailed discussion of acceptance and fair play, see Song, 2012, op. cit., pp. 212–15.

²⁰ Many objective list theorists will also hold that even objectively valuable goods are worth more or less to individual people depending on their particular projects. This complication matters less in the case of objective goods that are essential to any good life. Hedonism may be interpreted as a subjectivist view like desire-satisfaction theories, or as an objective list view with a short list. For a useful survey of views about wellbeing, see Kagan, S. (1998). *Normative Ethics*. Boulder, CO: Westview Press, pp. 29–41.

²¹ Dworkin, R. (2011). *Justice for Hedgehogs*. Cambridge, MA: Belknap Press of Harvard University Press, 2011, p. 303.

²² At worst, one might act viciously in failing to show gratitude in cases like that of the street musician. See Wellman, C.H. (1999). 'Gratitude as a Virtue,' *Pacific Philosophical Quarterly* 80(3), pp. 284–300. Another reason to be suspicious of the claim that one should generally reciprocate when one is benefited is that it may amount to double counting. After all, most

people manage to generate positive externalities of their own. See Schmitz, D. (2006). *Elements of Justice*. Cambridge: Cambridge University Press, pp. 90–92.

²³ David Lyons, for instance, characterizes fair play as being ‘for a just distribution’: Lyons, D. (1965). *Forms and Limits of Utilitarianism*. Oxford: Clarendon Press, p. 165.

²⁴ Cf. Simmons, A.J. (1979). *Moral Principles and Political Obligations*. Princeton, NJ: Princeton University Press, p. 77.

²⁵ Simmons, 2001, op. cit., p. 7.

²⁶ Simmons is surely right, however, that the argument Rawls provides for this requirement – that the injustice of the scheme undermines the voluntariness of the acceptance – is a non-sequitur.

²⁷ He may, however, demand a refund whether the contract generated rights or not. If it did, then he is owed reparation for a rights violation. If it did not, then the hitman has no claim to the payment he received.

²⁸ Simmons, 2001, op. cit., pp. 16–17.

²⁹ McDermott, D. (2004). ‘Fair-Play Obligations,’ *Political Studies* 52(2), pp. 220–22.

³⁰ Simmons writes elsewhere: ‘Frankly, I can see no good reason to insist that the enterprise be governed by rules. Mightn’t an enterprise be of the right sort which, say, assigned burdens fairly but not in accord with any preestablished rules? Cannot doing one’s part be obligatory under considerations of fair play even if “one’s part” is not specified by the rules?’ (2001, op. cit., p. 4). The problem that arises from McDermott’s objection if the rules condition is jettisoned is, perhaps, the good reason for which Simmons was looking.

³¹ In other words, I recognize that the wrongness of some cases of free-riding might depend on moral considerations other than fair play. Suppose, for instance, that (in some distant possible world) I enjoy public radio broadcasts, and listen to them frequently. Despite being told *ad nauseam* that such broadcasts are made possible by support from listeners like me, I do not take the hint, and never make any form of contribution to support the station. There are no rules about listeners contributing – the station may fear that any hint at the existence of rules, rather than looser social norms of reciprocity, might crowd out the most powerful motivations to contribute. Yet listening but refusing to contribute does seem like a clear case of free-riding. It is simply one that cannot be condemned on grounds of fair play. I am grateful to Colin Bird for pressing me on this case.

³² Zhu, J. (2015). ‘Fairness, Political Obligation, and the Justificatory Gap,’ *Journal of Moral Philosophy* 12(3), pp. 290–312. Calvin Normore also argues that the principle of fair play is a variant of consent. But Normore has in mind Peter Abelard’s psychological conception of consent, according to which ‘someone who was ready to perform an act in given circumstances, but for whom the circumstances never arose had consented to the act in the same sense as, and was morally on all fours with, someone for whom the circumstances did arise and who then performed it.’ Consent for Abelard (and Normore) is thus ‘an inward readiness to act should the circumstances arise.’ Normore, C. (2010). ‘Consent and the Principle of Fairness,’ in C. Favour, G.F. Gaus and J. Lamont (eds) *Essays on Politics, Philosophy, & Economics: Integration & Common Research Projects*. Stanford, CA: Stanford University Press. This rich and interesting view is, then, quite different from the modern ‘performative’ view of consent, which requires an observable sign of one’s will to agree.

³³ Zhu, 2015, op. cit., pp. 10–13.

³⁴ *Ibid.*, p. 12.

³⁵ Thus, despite Zhu’s claim, ordering in a restaurant is an example of express consent, not tacit consent. For a helpful account of the conditions for tacit consent, see Simmons, 1979, op. cit., pp. 80–83.

³⁶ Simmons: ‘An act may be such that it binds the actor morally to the same performance to which he would be bound if he had in fact consented. I may do something which is not itself an act of consent, but which nonetheless binds me as if I had consented’ (*ibid.*, p. 89).

³⁷ Zhu could insist that I have misunderstood the conventions of the open system in the public transportation case, but that would not save his argument, either. Even if he has in mind a

strange scenario in which skirting payment is a sign of consent to pay, there are still clear cases in which the acceptance of benefits does not signal consent. Suppose, for example, that Jones free-rides in some other public transportation system by hopping the turnstiles. In that case he accepts the benefits of the system while displaying an unambiguous sign that he does not consent to pay.

³⁸ Of course, one could also owe the same thing to different parties without conflict, and for precisely the same kind of moral reason. Suppose that Babe Ruth received not just one telegram asking him to hit a home run during the World Series for a sick child, but a dozen such telegrams. He could respond to signal his consent to every telegram, owe the home run to each child, and fulfill his duty by hitting one – assuming that no child greedily requested a home run for him- or herself alone.

³⁹ Gans, C. (1992). *Philosophical Anarchism and Political Disobedience*. New York: Cambridge University Press, p. 58n; Greenawalt, K. (1987). *Conflicts of Law and Morality*. New York: Oxford University Press, p. 140.

⁴⁰ Soper, P. (2002). *The Ethics of Deference: Learning from Law's Morals*. Cambridge: Cambridge University Press, p. 160.

⁴¹ Zhu, 2015, op. cit. I respond to Zhu specifically in Tosi, J. (2017). 'Playing Fair and Following the Rules,' *Journal of Moral Philosophy* 14(2), pp. 134–41.

⁴² Rawls, 'Legal Obligation and the Duty of Fair Play,' op. cit., p. 122.

⁴³ There must be some view of producer's rights at work here, but it is plausible enough on its face that absent other arrangements and special circumstances, people enjoy an ownership stake in what they produce. Since the productive agent in this case is a group, the group jointly owns the product of its joint labor. Rawls, for his part, thinks that this follows from 'the assumption of an initial position of equality.' *Ibid.*, p. 128.

⁴⁴ Surely Huemer must mean causally effective, rather than causally necessary. Otherwise there could be no obligation to support the provision of some goods, no matter how valuable they are, as there are sometimes no necessary costs – i.e. ones that could not be replaced were some other method chosen to produce the good.

⁴⁵ Huemer, M. (2013). *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey*. New York: Palgrave Macmillan, p. 84. Notice that Huemer's conditions are non-voluntarist. He claims that the lifeboat case (discussed below) suggests that acceptance is unnecessary. But the case shows no such thing. The case seems to rely on the intuition that it is wrong not to bail, perhaps to save oneself or others, or to best promote the good. It may even show that it is a wrong against your shipmates not to bail, depending on how one understands the duty to rescue. But it does not show that it is wrong not to bail because neglecting to do so is free-riding, or a violation of any other special obligation. Some further argument is needed to show that.

⁴⁶ *Ibid.*, p. 87. (emphasis in original).

⁴⁷ *Ibid.*, p. 164. (emphasis in original).

⁴⁸ I am grateful to audiences at New Mexico State University and the University of Virginia, where earlier drafts of this article were presented. For generous and helpful comments and discussion, I wish to thank Colin Bird, Thomas Christiano, Richard Dagger, Harrison Frye, Gerald Gaus, Dan Jacobson, Ross Mittiga, David Schmidtz, John Simmons, Steven Wall, and two anonymous referees.