

WILEY

Moral Free Riding

Author(s): Garrett Cullity

Source: *Philosophy & Public Affairs*, Vol. 24, No. 1 (Winter, 1995), pp. 3-34

Published by: Wiley

Stable URL: <https://www.jstor.org/stable/2265437>

Accessed: 28-08-2018 08:16 UTC

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/2265437?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Wiley is collaborating with JSTOR to digitize, preserve and extend access to *Philosophy & Public Affairs*

I. THREE KINDS OF FREE RIDING

“Free riding” is a term of art. To settle the correctness of an account of free riding, we must look not to an ordinary usage but to the interests in the service of which the notion has been invented. There are at least three such interests, and they generate distinct accounts.

What all three may be said to have in common is a conception of the free rider as someone who does not pay for goods that she consumes, where these goods are in some sense *public*.¹ Definitions of public goods vary widely, but they usually involve some subset of the following seven features:²

Jointness in Supply: if a public good is available to one member of the group for which it is public, then it is available to every other member at no cost to that other member.

Nonexcludability:³ if anyone is enjoying it, no one else (in the group for which it is public) can be prevented from doing so without excessive cost to the would-be excluders.

For helpful comments on this article, I am grateful to audiences at the Universities of Edinburgh and Glasgow, John Skorupski, Christopher Bryant, and the Editors of *Philosophy & Public Affairs*.

1. One sometimes encounters broader definitions, such as that of Donald Rutherford, *Dictionary of Economics* (London: Routledge, 1992), p. 181: “An individual who does not pay for the goods or services he/she consumes.” However, it is difficult to see what interest is served by adopting a definition as broad as this, which is satisfied by the recipient of a gift. (A similar formulation in David Gauthier, *Morals by Agreement* [Oxford: Clarendon Press, 1986], p. 96 is called a “definition” by his index.)

2. See the Appendix for a brief review of the public goods literature, and the provenance of my labels for the seven features.

3. Nonexcludability does not entail jointness in supply: consider deep-sea fish for those without boats. Furthermore, jointness in supply can be interpreted so as to include goods

Jointness in Consumption: one person's consumption of the good does not diminish the amount available for consumption by anyone else.

Nonrivalness: one person's enjoyment of the good does not diminish the benefits available to anyone else from its enjoyment.

Compulsoriness: if anyone receives the good, no one else can avoid doing so without excessive cost.

Equality: if anyone receives the good, everyone receives the same amount.

Indivisibility: there can be more than one consumer of the good, and each consumes the total output.

Under many conditions, the provision of public goods will be subject to the following theoretical problem. The failure of a group to cooperate to produce a public good may be collectively, but not individually, suboptimal from the point of view of self-interest: that is, there can exist a failure to cooperate, without any individual's acting suboptimally from that point of view, even though each individual is self-interested and is worse off without the good than he would be paying and getting it.⁴ A first interest in free riding is the game-theoretic one of stating the conditions under which this "free rider problem" arises, and the terms in which it may be solved.⁵ Notice, however, that the only kind of publicity required to produce this problem is *jointness in supply*—once the

that could be restricted, but are in fact made freely available (such as a pile of free newspapers), in which case it does not entail nonexcludability either. But if (as seems the usual practice) the conditional is read more strongly—as requiring that it be of the nature of any good exhibiting jointness in supply that its availability to one entails its availability to all—then nonexcludability does follow.

4. This does not make the free rider problem a Prisoner's Dilemma (contra Russell Hardin, "Collective Action as an Agreeable n-Prisoner's Dilemma," *Behavioural Science* 16 [1971]: 472–79): it does not yet say that for each individual, cooperation is suboptimal, irrespective of the actions of other group members. For examples that satisfy the description in the text without being Prisoner's Dilemmas, see Norman Frohlich, et al., "Individual Contributions to Collective Goods: Alternative Models," *Journal of Conflict Resolution* 19 (1975): 310–29; Jean Hampton, "Free-Rider Problems in the Production of Collective Goods," *Economics and Philosophy* 3 (1987); and Raimo Tuomela, "Free Riding and the Prisoner's Dilemma," *Journal of Philosophy* 85 (1988): 421–27. If you think that rational action is action that optimizes the agent's interests, then replace "suboptimal from the point of view of self-interest" with "irrational."

5. See Mancur Olson, *The Logic of Collective Action* (Cambridge: Harvard University Press, 1965); Frohlich, et al., "Collective Goods"; Richard Tuck, "Is there a Free-Rider Problem, and If So, What Is It?" in Ross Harrison, ed., *Rational Action* (Cambridge: Cambridge

good is produced, it can be acquired without paying—and not publicity in any of the other respects. In particular, the problem arises equally for rival goods, provided they are in joint supply, like limited stocks of fish in the sea. Given the game-theoretic interest, then, the best definition of a free rider is this: an individual who, in successfully optimizing his own interests, does not contribute to the production of a good that is in joint supply to a certain group, in conditions where it would be collectively self-interestedly suboptimal for the group not to cooperate toward its production.

To what extent is the theoretical problem a practical one? This depends, clearly enough, on the extent to which the conditions identified by the theoretical investigation are actually realized, and the extent to which actual agents approximate optimally self-interested ones. Here lies the interest of mainstream economics in the free rider: in particular, it wants to know how to construct mechanisms for the production of public goods to which beneficiaries have an incentive to contribute.⁶ Once we turn to tackling this practical problem, the other features of public goods (apart from jointness in supply) become important. In particular, the question whether the good is *rival* or not will have importance, given the effect on the motivation of many people of the knowledge that others stand to be harmed by their behavior. Thus for the economist's practical purposes, the best definition of a free rider is a different, looser one, employing a notion of publicity that ranges over the various

University Press, 1979); Philip Pettit, "Free Riders and Foul Dealers," *Journal of Philosophy* 83 (1986): 361–79; and Raimo Tuomela, "On the Structural Aspects of Collective Action and Free-Riding," *Theory and Decision* 32 (1992): 165–202. In calling this a game-theoretic interest, I am not suggesting that the definition of free riding is something one will find in a textbook of game theory.

6. See John O. Ledyard, "Incentive Compatibility," in John Eatwell, et al., eds., *The New Palgrave Dictionary of Economics* (London: Macmillan, 1987), p. 739; Tyler Cowen, ed., *Public Goods and Market Failures* (New Brunswick: Transaction, 1992); Jack Hershleifer, "From Weakest Link to Best-Shot: The Voluntary Provision of Public Goods," *Public Choice* 41 (1983): 371–86; and R. Mark Isaac, et al., "Public Goods Provision in an Experimental Environment," *Journal of Public Economics* 26 (1985): 51–74. The economist's problem ramifies into the work of political scientists such as Norman Frohlich and Joe A. Oppenheimer, "I Get by With a Little Help From My Friends," *World Politics* 22 (1970): 104–21; idem. and Oran R. Young, *Political Leadership and Collective Goods* (Princeton: Princeton University Press, 1971); and Frohlich and Oppenheimer, *Modern Political Economy* (Englewood Cliffs: Prentice-Hall, 1978); experimental sociologists such as G. Marwell and Ruth E. Ames, "Experiments on the Provision of Public Goods, I," *American Journal of Sociology* 84 (1979): 1335–60; and social psychologists such as D. M. Messick and M. B. Brewer, "Solving Social Dilemmas: A Review," *Psychological Review Papers* 11 (1983): 44.

features identified above: a free rider is a person who fails to pay for a public good she consumes, although the good is worth paying for.⁷

But a third, and distinct, interest in free riding is the focus of this article. This is the distinctively moral philosophical interest in examining the conditions under which a certain sort of *unfairness* is commonly thought to arise—an unfairness of which the paradigm is:

Fare-Evasion

Public transport in my town is efficiently run on an “honor” system, which places the onus on passengers to buy a ticket before traveling and to cancel it in a machine on any vehicle they use. I ride without paying.

Given this interest, producing an account of free riding means specifying the conditions under which unfairness of this sort arises, and explaining why those conditions yield unfairness. Producing such an account is my aim in what follows.

Any account of free riding of this third sort must be distinct from the two already stated, for the simple reason that on either of the previous definitions, free riding need not be unfair. (The usual explanation of this is that the “free ride” can be taken on a scheme that is itself unfair or otherwise objectionable, such as a cartel.⁸) Moreover, it is significant that, once more, free riding in this third sense will not require the presence of all of the features of publicity identified above. Someone who sneaks into a private theater without paying would appear to be exhibiting the same sort of unfairness as the Fare-Evader, although the goods the sneak enjoys are public only in the sense of being *nonrival*: in enjoying them, he does not harm anyone else. And there is a reason why the moral philosopher should take a special interest in this feature of cases of unfairness. For an important stimulus to our thinking about fairness is provided by the following question: “If the free rider harms no one, what is it about her conduct that makes it unfair?” A moral philosophical account of free riding, as I shall understand it here, offers an answer to this question. The definition it yields will accordingly have the following

7. Compare David W. Pearce, ed., *The Macmillan Dictionary of Modern Economics* (London: Macmillan, 1981), p. 165.

8. See E. C. Pasour, “The Free Rider as a Basis for Government Intervention,” *Journal of Libertarian Studies* 5 (1981): 453–64. As Frank Miller and Rolf Sartorius, “Population Policy and Public Goods,” *Philosophy & Public Affairs* 8, no. 2 (Winter 1979): 152, point out, many of the commodities meeting standard definitions of public goods are positive evils.

schematic form: a free rider is someone whose failure to pay for nonrival goods under conditions C makes her conduct unfair.

I said that unfairness “is commonly thought to arise” in cases such as Fare-Evasion. I do not wish to deny that there are conceptions of fairness according to which Fare-Evasion, sneaking into theaters and the like are not unfair, either because such actions harm no one or for some other reason.⁹ The ambition of a moral philosophical account of free riding, as I conceive it, will not include showing that one *must* conceive of free riding as unfair—that the failure to do so shows that one is conceptually confused about fairness. The ambition is rather to show what, on *this* conception of fairness, counts as unfair, and why it should be thought to do so. Later, I shall offer an explanation of what unifies this conception, and why “unfairness” should be thought an appropriate term for the cases unified under it. However, an adjudication between this and other, rival conceptions of fairness (if it can be performed) is not performed here. The claim for which I shall be arguing is contentious enough as it stands: if Fare-Evasion is unfair, then so are some refusals to pay for benefits nonvoluntarily conferred upon one.

The two tasks to be pursued here are to specify the conditions C under which the failure to pay for nonrival goods is unfair, and to show *how* the satisfaction of the specified conditions contributes to making someone’s conduct unfair. Now several writers have claimed to complete at least the first of these tasks, in advocating versions of what is often called the “Principle of Fairness.” John Rawls first used the label in *A Theory of Justice* to refer to the following claim:

a person is [morally] 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), that is, it satisfies the two principles of justice; and second, one has voluntarily accepted the benefits of the arrangement or taken advantage of the opportunities it offers to further one’s interests.¹⁰

The ensuing literature has been governed by an interest in whether a plausible refinement of this principle will provide a foundation for polit-

9. That is not to say that proponents of such conceptions must deny that free riding is wrong. Their most natural alternative attempts to account for its wrongness as a matter of theft, and to explain the difference between theft and the permissible taking of goods without appealing to fairness.

10. John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1971), pp. 111–12. The debt to Hart’s “principle of mutual restriction” is acknowledged when Rawls adds

ical obligation;¹¹ and to this end, the central issue of contention has concerned the relaxation of Rawls's requirement that the benefits in question be *voluntarily accepted*. My own formulation of the Principle of Fairness, presented intuitively in Section III, endorses this relaxation, while differing extensionally from other such proposals. More significantly, it is defensible on grounds more substantial than conformity with "intuitive" (i.e., pretheoretical) judgments concerning imagined examples, and *ad hominem* arguments against the positive proposals of its opponents. For in Section IV, I address the second task mentioned above, of showing *how* the features contained in the principle contribute to making a person's conduct unfair.

But before redeeming these claims, let me begin with a quick look at some of the most prominent work on the same question.

II. VOLUNTARY AND NONVOLUNTARY BENEFITS

According to Rawls's stated Principle of Fairness,¹² there are two features in virtue of which the conduct of someone like the Fare-Evader is

the remark: "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." Compare H.L.A. Hart, "Are There Any Natural Rights?" *Philosophical Review* 64 (1955): 185; and for an important antecedent to Hart, see C. D. Broad, "On the Function of False Hypotheses in Ethics," *International Journal of Ethics* 26 (1916): 384–90.

11. See A. John Simmons, "The Principle of Fair Play," *Philosophy & Public Affairs* 8, no. 4 (Summer 1979), and idem., *Moral Principles and Political Obligations* (Princeton: Princeton University Press, 1979); Richard J. Arneson, "The Principle of Fairness and Free-Rider Problems," *Ethics* 92 (1982): 618–19; George Klosko, "The Principle of Fairness and Political Obligation," *Ethics* 97 (1987), and idem., *The Principle of Fairness and Political Obligation* (Lanham: Rowman and Littlefield, 1992); Vincent Maphai, "The Principle of Fairness," *South African Journal of Philosophy* 6 (1987): 73–80; and Michael Davis, "Nozick's Argument for the Legitimacy of the Welfare State," *Ethics* 97 (1987): 576–94. It is rarely observed that providing such a foundation is not the intention in Rawls, *Theory*, which abandons the advocacy in John Rawls, "Legal Obligation and the Duty of Fair Play," in S. Hook, ed., *Law and Philosophy* (New York: New York University Press, 1964) of a similar "Principle of Fair Play" for that end. The later claim is that the Principle of Fairness grounds all obligations, including most prominently the obligation of fidelity—i.e., promise-keeping—but that political duties have a different source.

12. His stated principle, as opposed to the ostensible clarification expounding its "main idea," which appears to add the further conditions that the institution in question must be cooperative, and must be mutually advantageous. See note 10.

unfair: the first concerns the justice of the benefit-conferring institution, and the second the voluntary acceptance of its benefits.

Now of course, Rawls's claim is only that these features are jointly sufficient for unfairness: he is not denying that there are other sorts of unfairness as well. However, A. J. Simmons was surely right to complain that his inclusion of the first, "justice condition" results in an unduly restricted claim. For if I have already voluntarily accepted the benefits offered by an institution, surely I cannot *then* start raising objections about its fairness, with a view to justifying nonpayment. Even if my complaint is that the institution treats *me* unfairly by demanding a disproportionate share of the cost, I cannot claim this, having voluntarily accepted its benefits, as a justification for not paying altogether. Indeed, the same point would appear equally to militate against Simmons's own preferred version of the principle—that if one has voluntarily accepted the benefits conferred by a cooperative scheme, then one is obligated to contribute one's fair share of the costs of its production.¹³ For if I have voluntarily accepted the benefits, knowing the associated cost, it still seems too late to start caviling about my fair share. If I had a complaint about the amount being asked of me, I ought not to have taken the benefits in the first place.

However, the more important and controversial question concerns Rawls's second condition, according to which one must have "voluntarily accepted the benefits of the arrangement or taken advantage of the opportunities it offers to further one's interests." It is not clear exactly how Rawls intends this to be construed: how inclusive a reading should be given to the latter clause? In common with others,¹⁴ I shall maintain that unless it is given an extremely inclusive reading, this condition is also too restrictive. We should endorse what I shall call an *extended* Principle of Fairness—one according to which it can be unfair to refuse to contribute to the production of benefits that have been conferred upon one nonvoluntarily. However, anyone advocating such a view

13. See Simmons, "Principle of Fair Play," p. 317. For further apparent counterexamples, see Arneson, "Principle of Fairness," p. 620.

14. See Nora K. Bell, "Nozick and the Principle of Fairness," *Social Theory and Practice* 5 (1978): 65–73; Arneson, "Principle of Fairness"; Christopher Morris, "The Hart-Rawls Principle of Fairness Amended," *Journal of Social Philosophy* 14 (1983): 18–20; Maphai, "Principle of Fairness"; Davis, "Nozick's Argument"; and Klosko, "Principle of Fairness" and *Principle of Fairness*.

confronts the vigorous and influential opposition to such a relaxation expressed by Robert Nozick in a few pages of *Anarchy, State and Utopia*. As Nozick's lively examples suggest, surely there is something morally objectionable about the idea that you might be able to obligate others to pay you by pressing benefits on them against their will: this would appear to ground an accusation of unfairness against the benefactors rather than the beneficiary. An example of my own gives the type to which Nozick's belong:

The Enterprising Elves

On the first day in my newly carpeted house, I leave my shoes outside. In the morning I am delighted to find they have been extraordinarily well repaired. I am less delighted when I receive the bill.¹⁵

Here, it seems that all I need to say to support the fairness of refusing to pay is that the benefits were forced on me. To this basic point Nozick adds a number of subsidiary ones against possible schemes for conferring nonvoluntary benefits: if the benefits are not worth the demanded cost,¹⁶ or they are not fairly distributed,¹⁷ or my conscientious objection to the scheme is justified,¹⁸ then this only adds to the case for the fairness of refusing to pay. However, the main point would seem to be independent of these further ones. In the case just imagined, my refusing to pay would be fair, even if the benefit-conferring institution possesses none of these defects—the work is worth the money, costs and benefits are distributed fairly, and there are no grounds for conscientiously refusing to support this institution in favor of an alternative: it is enough, it seems, that the benefits were unsolicited.

15. Compare Robert Nozick, *Anarchy, State and Utopia* (New York: Basic Books, 1974), p. 95, for the nuisance who throws books into people's houses and then demands payment. In aiming his discussion against Rawls, Nozick assumes the inclusive reading of his principle. Simmons, "Principle of Fair Play," defends Rawls against the criticism by offering a stronger reading, which stresses voluntary acceptance, with the gloss (p. 327) that this involves one's "either (1) trying to get (and succeeding in getting) the benefit, or (2) taking the benefit willingly and knowingly." For a more detailed discussion, see Simmons, *Moral Principles*, pp. 42–45, 89–90, 138–63, 187–99. See also Miller and Sartorius, "Population Policy," p. 166.

16. See Nozick, *Anarchy*, pp. 93–94, for the well-known example of the public address system.

17. See *ibid.*, p. 94.

18. See *ibid.*, pp. 94–95.

This judgment may seem persuasive, but so also, to several writers, does the judgment that the refusal to pay is *not* fair in a case such as the following:

The Recalcitrant Fisherman

Pollution from the boats fishing our lake has become serious enough to affect the catch. Every fisherman agrees to stop polluting the lake, and contribute toward cleaning it up—except one, who protests, “I have not chosen to receive these benefits, nor have I misled you into conferring them on me. I am simply going on as I always have done. If you do not want to benefit me find another lake.”¹⁹

If the community is large enough, perhaps his practices are not detrimental to other people (his pollution alone may not affect the fish). And if we suppose that the others’ actions enhance the health of the fish, but not their number, then it is hard to see how he *accepts*, in any substantial sense, the benefits conferred upon him.²⁰ I submit that he is still being unfair.

The example, of course, is bound to be controversial. However, a line of objection that should *not* be maintained against it is that since non-payment for unsolicited benefits is fair in the earlier, Nozickian example of the Enterprising Elves, it must be equally fair here.²¹ For not all cases of nonpayment for unsolicited benefits are the same. If we follow Richard Arneson, the key to reconciling the Nozickian resistance to paying in the former sort of case with a complaint of unfairness against non-payment in the latter, is to observe that the former case deals in *exclud-*

19. The lighthouse example of Gauthier, *Morals by Agreement*, p. 96, is an analogue of this one.

20. If he were catching more fish thanks to their self-restraint, then it could be maintained that in choosing not to throw back the surplus, he would be accepting the benefits conferred upon him.

21. This line of thought seems implicit in Nozick, *Anarchy*, and is explicit in Simmons, “Principle of Fair Play.” Simmons is prepared to accept that “ordinary feelings about fair play” do endorse an accusation of unfairness in this sort of case (in his example, a resident enjoys the benefits of a cooperative scheme for beautifying the neighborhood, while refusing to join in—pp. 330–31), but maintains that “those feelings are mistaken” (p. 332). However, it is only by reference to such “ordinary feelings” that his own Principle of Fairness has been supported. If, therefore, a more discriminating version of the principle can be provided, accommodating the Nozickian judgments with which Simmons begins, while also endorsing the accusation of unfairness in this sort of case, then Simmons’s methodology would appear to compel him to accept it.

able goods, while in the latter the goods are nonexcludable.²² However, it is not as if it is open simply to claim that, if I receive nonexcludable benefits from a scheme of any kind, it is unfair to refuse to meet its requirements of payment. The six further conditions that Arneson thinks must be added are assembled by him into the following version of the Principle of Fairness:

where a scheme of cooperation is established that supplies a collective [i.e., nonexcludable] benefit that is worth its cost to each recipient, where the burdens of cooperation are fairly divided, where it is unfeasible to attract voluntary compliance to the scheme via supplementary private benefits,²³ and where the collective benefit is either voluntarily accepted or such that voluntary acceptance of it is impossible, those who contribute their assigned fair share of the costs of the scheme have a right, against the remaining beneficiaries, that they should also pay their fair share. (P. 623)

It seems to me that there remain problems with the extension of this proposal, and likewise for those more recent proposals made in the same spirit. But on this score, let me simply note that, since the benefits received in the original Fare-Evasion example are excludable, its unfairness is not entailed by Arneson's principle—so his principle can hardly serve to supply an account of free riding of the kind I have set out to produce.²⁴

A more fundamental point is this: how can a proposal like this one

22. See Arneson, "Principle of Fairness," pp. 617–20; also Klosko, "Principle of Fairness," p. 353–55, and Maphai, "Principle of Fairness," pp. 77–78.

23. Arneson, "Principle of Fairness," p. 622, spells this condition out more fully: "It is unfeasible that the cooperative scheme be arranged so that private benefits are supplied to each beneficiary of B in sufficient quantity to induce all beneficiaries to contribute their fair share of the costs of the scheme." The rationale of this condition "is simply that, if one can secure the needed public good in a fair manner and without coercion, one should not resort to coercion."

24. It might be held that although the benefits conferred by other public transport schemes are excludable, those conferred by this one, with its honesty system, must be nonexcludable. However, this move, which would make excludability more properly a characteristic of benefit-conferring schemes than of the benefits themselves, has the more serious consequence that goods of any type conferred by a scheme an essential characteristic of which is its forcing them upon people will count as nonexcludable. This would mean endorsing the demands of a cooperatively organized version of the Enterprising Elves' scheme.

Anyway, this maneuver does not address the theater-sneak case, which equally exemplifies the sort of free riding of which I am seeking an account.

claim to supply a full answer to the question with which I began? That is, given the conditions under which it claims a person's conduct to be unfair, *how* does the satisfaction of those conditions contribute to making it unfair? The force of this question is independent of any concerns one may have about the extension of the principle. For even if one were disposed to accept all the judgments falling under it, one is still entitled to question whether it is the satisfaction of its conditions that *makes* a person's conduct unfair. Of course, it is not as if this sort of question will always be intelligible. To ask how its expressing a delight in the suffering of others contributes to making an action cruel, for instance, is to invite the response that one cannot have grasped the concept of cruelty. However, the connection between unfairness and the various conditions of the principle just stated can hardly be claimed to exhibit this degree of transparency.

Moreover, the force of the question is surely intensified by the existence of determined opposition to the extension of the Principle of Fairness to cover nonvoluntary benefits. A thoroughgoing opponent will simply fail to find any plausibility in intuitive judgments such as the one advanced above about the Recalcitrant Fisherman.

Now it is not obvious what, if anything, there is to be said against this sort of opponent. Certainly, other proponents of extended versions of the principle are not sensibly thought of as seeking to address him.²⁵ They clearly start from premises by which he need not be impressed, and it is not obvious that we are entitled to expect any better. But we can in fact do better—or so I shall argue. After expounding my own reformulation of the principle, I shall offer an argument in its support that takes on directly the sort of thoroughgoing opponent just described, against whom attempts at intuitive illustrations of its plausibility make no headway. From premises even he should accept, an extended Principle of Fairness can be derived.

25. There is surely no threat to such an opponent in either the observation that the judgments Nozick actually makes are compatible with recognizing obligations to pay for nonvoluntary benefits in other circumstances (compare Maphai, "Principle of Fairness," p. 74); the strategy of asserting the evident nature of political obligations and claiming a version of the Principle of Fairness as their best explanation (Klosko, "Principle of Fairness," p. 355); or the interesting argument that Nozick's Lockean account of the legitimacy of the appropriation of private property in a state of nature relies upon foundations that equally support an extended Principle of Fairness (Arneson, "Principle of Fairness," pp. 623–33; compare Davis, "Nozick's Argument").

III. A PRINCIPLE OF FAIRNESS

The problem faced by anyone who would defend an extended Principle of Fairness (according to which refusing to meet demands to pay for nonvoluntary benefits can be unfair) is that of discriminating circumstances where such demands are reasonable from those where they are merely predatory. We have seen Arneson's proposal for doing so. But now consider this simpler suggestion, which is a prototype of the principle to be advocated here.

If a person is benefited by a scheme²⁶ that makes fairly distributed requirements, the benefit is worth its cost, and it is not the case that practically everyone would be made worse off by the practice of regarding as obligatory²⁷ those further requirements that must in fairness be regarded as obligatory if the requirements in question are regarded as obligatory, she is being unfair.

To see how this precludes the predatory demands, recall the earlier example. Having had my shoes repaired, I have benefited from the Enterprising Elves' shoe-repairing scheme; and since they do not overcharge for their services, the requirements of their scheme are fairly distributed and the benefits worth their cost. However, consider what would be entailed if we were to recognize as obligatory all the further demands that would in fairness have to be so regarded if the demands of the Enterprising Elves were regarded as obligatory. (I shall call this *fairly generalizing* the demands made by the Enterprising Elves.) It would mean holding everyone liable to pay for all unsolicited benefits that are worth their cost. A commercial system that recognized this sort of liability would be so cripplingly inefficient that it would impoverish us: it is clearly better for practically everyone if commercial transactions can only be entered into by means of an explicit act of commitment. By requiring, therefore, that the fair generalization of the scheme's requirements must not make practically everyone worse off, my principle accommodates the right conclusion, that my not paying my benefactors

26. The deliberately vague term "scheme" is used here to range over any actual or potential practice that confers benefits by making requirements of beneficiaries.

27. I help myself here to a loose usage of "obligatory," according to which to regard a requirement as obligatory is simply to regard it as a requirement that it would be morally wrong not to comply with. For a more discriminating usage, see e.g. E. J. Lemmon, "Moral Dilemmas," *Philosophical Review* 71 (1962): 141; and Rawls, *Theory*, p. 113.

in this case is fair. In contrast, the fair generalization of the requirement imposed by the clean fishing scheme that benefits the Recalcitrant Fisherman is the making of a fairly assessed sacrifice by each in order to preserve the livelihood of all. *This* would be beneficial to each person of whom the requirements are made, and detrimental to no one else.

A word on “practically everyone.” It would be too strong to require that *no one* is to be made worse off by the fair generalization of the scheme’s requirements: there might be a few resourceful types (perhaps the Elves themselves) who would stand to benefit in the chaos resulting from the shoe-repairing scheme’s fair generalization, but *this* surely does not militate against the fairness of my nonpayment. In accommodating this point, though, by requiring only that *practically* everyone is not to be made worse off by the scheme’s fair generalization, I may have raised a different worry. I might seem to be implying that it is always fair to promote benefits to the overwhelming majority at the expense of a minority, which looks false. This worry is met by stipulating: “practically everyone” means either everyone, or enough others, and in such circumstances, as to make it unreasonable for the minority to resist the outcome that is better for the majority.²⁸

Now most previous formulations of the Principle of Fairness can disqualify demands of the sort exemplified by the Enterprising Elves on at least one of two grounds: the benefits in question are excludable, and the scheme that confers them is not cooperative.²⁹ However, the version of the principle just stated includes neither of these conditions. That this is a point in its favor is suggested by:

The Shoe-Repairing Convention

I move into an area where there is a well-established convention (of

28. According to Nozickians, if an outcome is produced in circumstances where payment is required for nonvoluntary benefits, then any minority may reasonably resist it. But that should not persuade you that they can agree with my principle. Their claim is that it is not unreasonable, under such circumstances, for a minority to resist; so they are not claiming that practically everyone (in the stipulated sense) is worse off. If so, the principle does produce anti-Nozickian judgments concerning the unfairness of nonpayment for nonvoluntary benefits, even if we allow the Nozickians’ own claims concerning reasonableness to govern the interpretation of “practically everyone.”

29. For formulations that include a cooperation condition, see: Hart, “Are There Any Natural Rights?” p. 185; Rawls, *Theory*, p. 112; Simmons, “Principle of Fair Play,” p. 317; and Morris, “The Hart-Rawls Principle,” p. 18. For both cooperation and nonexcludability conditions, see: Arneson, “Principle of Fairness,” p. 623; Klosko, “Principle of Fairness,” pp. 353–55; and Maphai, “Principle of Fairness,” p. 78.

which I am unaware) that leaving one's shoes outside amounts to a request to have them fixed. But when mine are repaired, I refuse to pay.

This scheme is not cooperative (except in the sense in which the maintenance of any convention is), and it confers excludable benefits. But this time, my refusal would be unfair: in particular, my complaint that the cleaning of my own shoes was unsolicited would appear to be beside the point. For unless there is a reason for holding that the shoe-repairers' practice itself violates a commercial code of overriding validity,³⁰ the fact that there is a fair (if unusual) convention governing commercial transactions in the community I have chosen to join vitiates my complaint. Notice, moreover, that it does not seem to matter whether I knew of their practice before moving or not. If someone brought up in a culture where it is acceptable to taste food before committing oneself to a purchase does so on his first visit to a British or American grocer's, there may be no question of blaming him for doing so, but the shopkeeper is entitled to demand payment all the same.

The unfairness of my refusing to meet the demands of the Shoe-Repairing Convention is entailed by the principle stated above. For here, unlike the case of the Enterprising Elves, people can choose which benefits to receive, by choosing whether to leave their shoes out or not. No chaotic perils lie in the adoption of their alternative commercial convention, so the fair generalization of the scheme remains acceptable. And the unfairness of the Fare-Evader and the theater sneak with whom I began are similarly accommodated.

I promised a defense of my principle that goes beyond this sort of appeal to consonance with intuitive judgments. Before providing it, however, the principle requires some modification.

In particular, let us examine the clause requiring that the benefit received must be worth its cost. The need for greater precision here is most clearly displayed in the case of benefits that are luxuries. Suppose I am compulsorily benefited by a new scheme for cleaning and maintaining the pavements to an exceptionally high standard; and that the work is very efficiently carried out, so that the bill for my contribution, although substantial, is not exorbitant; but that I am quite poor and do not particularly care whether the pavements are immaculate rather than

30. As would be the case if they had simply congregated in a neighborhood of St. Andrews and begun trying to impose their commercial practice upon others.

merely tolerable. Is the benefit worth the cost? There is clearly *a* sense in which it is—I am not being overcharged for the service—but surely more relevant is the sense in which it is not: all things considered, I would be *worse* off getting the benefit and paying than if I did neither.

However, this latter suggestion requires amplification in three further respects. First, not every benefit conferred by a scheme is relevant to drawing conclusions about the unfairness of nonpayment. If a piece of treasure turns up in the belly of one of the newly energetic fish caught by the Recalcitrant Fisherman, *this* consequence of others' participation in the fishing scheme is surely irrelevant to any complaint of unfairness against him. Why do we regard this kind of benefit as extraneous? Evidently, because of the description under which the scheme's benefits are produced. The aim toward which the fishermen are cooperating is not the discovery of treasure but the protection of their livelihoods, so only the latter benefits are relevant to an accusation of unfairness.

A similar question arises in relation to the costs imposed by a scheme. Which of these, in addition to those explicitly designated as payments for the scheme's benefits, bear on the fairness of nonparticipation? Some evidently do: if the Recalcitrant Fisherman can observe that the process of rendering the lake safe for fish has made the water unsafe for his family to swim in and drink, surely this lends support to his complaint that the demand on him for payment is unreasonable. But some costs are clearly irrelevant: if the Fare-Evader is eventually run over by a bus, this does not show that her fare-evasion was justifiable after all. What explains the difference? The general form of the explanation seems to be this. A cost that an individual suffers as a consequence of others' participation in a benefit-conferring scheme bears on the fairness of his nonparticipation when the others, in participating, are collectively *morally answerable* for those costs—where this means that they ought, morally, either to be able to supply a morally good reason for producing them, or to compensate those suffering them.³¹

The third complication is this. When determining the sense in which one must receive a net benefit from a scheme (for noncompliers to be appropriately accused of unfairness), we need to *exclude* from consider-

31. I need not try to say here in virtue of what one agent is to be held morally answerable for imposing a cost on another. My claim requires only that our difficulties in deciding what an agent is morally answerable for are themselves difficulties concerning the attribution of costs to people's participation in a benefit-conferring scheme.

ation any benefits that result from the scheme, but which depend on others' *non*participation. This point applies to contexts where there is partial compliance with a heavily demanding scheme. Suppose the fish in our example are being poisoned by a substance all the fishermen are using, and that when several refrain from using it, the fish return to perfect health. But suppose also that without the regular introduction of moderate quantities of this substance into the lake, the fish would all have died. Now consider the proposal that I refrain from using the substance altogether. I have benefited as a result of others' participation in this scheme, and perhaps the cost of compliance does not outweigh that benefit. But if the scheme requires me not simply to moderate the amount of the substance I use but to refrain from using it altogether, then the following point is significant: the scheme only produces its benefit by relying on others' noncompliance with this requirement. (Without the noncompliance all the fish would be dead.) So against the demand that I comply with this scheme's requirement, there seems to be an effective reply: I'm actually benefited more by the noncompliers than by the compliers. And if so, then the accusation that I am unfairly free riding on the latter is surely undermined.

The foregoing remarks suggest the following interpretation of the requirement that the benefit received by a free rider must be worth its cost. It must be the case that the balance of the benefits and costs attributable (in the specified sense) to others' participation in the scheme, plus the cost of my own compliance with the scheme's requirements, but excluding from consideration any benefits attributable (in that sense) to others' *non*participation, is positive. Let us call this the requirement that the *practice of participation* in the scheme is to represent a net benefit for me.

The final version of my Principle of Fairness can now be stated.

If a person receives benefits from a scheme that satisfies the following conditions, it is unfair for her not to meet the requirements it makes of her in respect of her enjoyment of those benefits.

- (i) The practice of participation in the scheme represents a net benefit for her.
- (ii) It is not the case that practically everyone would be made worse off by the practice of participation in the recognition as obligatory of

those further requirements that must in fairness be regarded as obligatory if the requirements in question are regarded as obligatory.

(iii) She is not raising a legitimate moral objection to the scheme.

Condition (i) elaborates on the earlier formulation of the principle in the manner just explained—replacing its clause requiring that the benefits in question be worth their cost with the requirement that the practice of participation in the scheme represent a net benefit. Condition (ii) qualifies, in the same way, the clause requiring that the “fair generalization” of the scheme’s requirements not be to practically everyone’s detriment. And condition (iii) reintroduces Rawls’s requirement that the benefit-conferring scheme be a just one, modifying it in two respects. First, it concerns moral objections to a scheme beyond those of injustice: the objection might be not that there is anything unjust about the scheme, but that it is cruel, wasteful, or degrading.³² But secondly, it requires that the act of noncompliance must itself be the raising of a legitimate moral objection in order to be fair. For it seems intuitively that gangsters can free-ride on their extortion rackets, too: their unfairness to their associates is not dissolved by the moral objections to the schemes themselves.

Now it may seem that, despite these modifications, the third condition faces the problem Simmons raised for Rawls’s justice condition. My principle purports to apply both to schemes that confer benefits nonvoluntarily and to those that make benefits available for voluntary acceptance. But in the latter case, surely Simmons is right to claim that my *acceptance* of the benefits vitiates any attempt to justify nonpayment by complaining that the institution is morally objectionable. Suppose I join a time-sharing holiday scheme: I cannot take my holiday and then refuse to pay for it on the grounds that other members of the scheme are being overcharged. And if the principle cannot accommodate this judgment, then that seems to be a serious defect—for the unfairness here seems to be of the same paradigmatic sort exhibited by the Fare-Evader, which the principle set out to capture.

But *is* such a judgment incompatible with my principle? Let us examine it more closely. If we judge nonpayment in such a case to be objec-

32. Of course, we are still at this point awaiting my argument against the Nozickian claim that a scheme’s requiring someone to pay for benefits it has forced upon him itself grounds a legitimate moral objection.

tionable, it is because we endorse the convention that accepting benefits at an accessibly advertised price binds one to pay that price. But having noted the existence of this convention, the elasticity of the notion of a “scheme” employed by the principle becomes significant. For the convention just described clearly qualifies as a benefit-conferring scheme in my sense. Given the existence of first-order benefits such as holidays, we can consider the following second-order benefit: the opportunity to enjoy such first-order benefits if one wishes. In respect of this benefit, a requirement is conventionally made—the requirement that one restrict oneself in the following conditional way: *if* one accepts first-order benefits that are made available at an accessibly advertised price, one is bound to pay that price.

Not only does this convention qualify as a scheme in my sense; it also satisfies the principle’s three conditions. The practice of participation in this scheme represents a net benefit, since the benefit conferred by others’ participation (the opportunity to acquire first-order goods) is worth its cost (being restricted to acquiring them only if one pays). Its fair generalization is not to practically everyone’s detriment. And there appears to be nothing morally objectionable about the convention: the fact that its benefits and requirements are conferred equally upon everyone seems to make it free of distributive unfairness, and there appears to be no other moral flaw. If so, the unfairness of refusing to pay for the voluntarily accepted benefits of an immoral scheme can be explained as the unfairness of noncompliance with this second-order scheme. Simmons’s point, it transpires, is a corollary of my principle.³³

Notice that the proposed principle really is a principle of *fairness*. It is concerned to specify the conditions under which nonpayment is unfair, rather than to say anything about any rights or entitlements that the producers of a benefit may have. In this, it differs from previous proposals; and more particularly, it does not as it stands claim to set out conditions for the legitimacy of *coercing* beneficiaries to contribute to a benefit-conferring scheme.

This modesty admittedly reduces the principle’s interest, since it evidently precludes drawing the sorts of conclusions concerning political obligation that others have sought from a Principle of Fairness. How-

33. The same strategy of argument will support obligations to pay for benefits that have been accepted from a scheme failing to satisfy condition (i) or (ii)—for instance, one whose costs outweigh its benefits.

ever, in reaching for these conclusions, the other accounts commit themselves to unacceptably strong claims. My own claim concerning the Recalcitrant Fisherman has been that his noncompliance is unfair; according to the previous attempts, the compliers can *force* him to pay. What is wrong with the stronger claim? It is that if there is any question of the use of force, it should be applied by the police, rather than those making the complaint. Why? For the reason captured in condition (ii) above. If we accepted the entitlement of the fishermen here to enforce their own complaints of unfairness, we must in fairness extend a similar entitlement to everyone else. However, this would be worse for practically everyone than ceding such powers of enforcement to the police. And this gives the objection its final twist. In enforcing their own complaints of unfairness, the fishermen would be arrogating privileges to themselves while relying (for the benefits conferred upon them by the rule of law) upon the renunciation of those privileges by others. Their complaint of unfairness would rebound upon them.

But if this is right, it suggests that a less reticent, but closely related, principle governs the case where a benefit-conferring scheme is coercive:

If a person receives benefits from a *coercive* scheme that satisfies conditions (i)–(iii), but does not meet the requirements it makes of her in respect of her enjoyment of those benefits, the scheme’s prescribed methods of coercion may fairly be used to compel her to meet them.

If they were seeking themselves to employ coercion, then the fishermen’s scheme would fail to satisfy condition (ii), since its fair generalization would be detrimental to practically everyone.³⁴ However, the unfairness of failing to contribute to their noncoercive scheme will be endorsed by the primary version of the principle.

The principle governing coercive schemes gives us a plausible ingredient in an argument for political obligation. However, it cannot be part of the ambition of the present paper to complete such an argument. For by far the larger part of that task would be the defense of the claim, for any given scheme of political coercion, that the practice of participation in it does indeed represent a net benefit for a putatively obligated citi-

34. Notice that this goes beyond Arneson’s sensible suggestion that we take the costs of coercion into account in assessing whether a scheme confers a net benefit (Arneson, “Principle of Fairness,” p. 621): even so, its fair generalization may fail to be beneficial.

zen. Since this would mean showing that the scheme's constraints on personal autonomy are outweighed by the benefits it produces, it is hard to see how such a claim could be defended without first producing an account of the value of autonomy.

A moral philosophical account of free riding, I observed at the outset, is one governed by an interest in answering, in cases like Fare-Evasion, the question, "If the free rider harms no one, what is it about her conduct that makes it unfair?" A moral philosophical definition of the free rider, I remarked, would therefore have the schematic form: a free rider is someone whose failure to pay for nonrival goods under conditions C makes her conduct unfair. The version of the Principle of Fairness that has now been expounded and given intuitive support specifies those conditions. This completes my first task. But there remains a second: to explain *how* the satisfaction of the specified conditions contributes to making nonpayment unfair. And I have undertaken to offer that explanation to an opponent who fails to find any intuitive plausibility in an extended Principle of Fairness.

IV. JUSTIFICATION OF THE PRINCIPLE

Nothing short of a fully fledged theory of justice would count as *completely* fulfilling this second task. However, it is possible to make progress here without anything so ambitious. My argument starts simply by assuming the unfairness of Fare-Evasion—not that this judgment deserves any especially privileged status; but it does appear to be a reasonable starting-point, and it gains support to the extent that the following discussion can integrate it into a recognizable conception of fairness. Anyone who accepts this, I shall now argue, should accept the extended Principle of Fairness formulated above.

The Fare-Evader shows us something important about fairness—that particular acts of unfairness need not harm anyone. If not, where does the unfairness in her conduct lie? An answer of the most general and uncontentious kind is this: the Fare-Evader's unfairness is a matter of her giving herself objectionably preferential treatment. In seeking the benefits she does, the Fare-Evader depends upon the willingness of others to subject themselves to a requirement to pay, without being willing to do so herself. The benefits only exist because others who seek them take it upon themselves to contribute toward their production: in taking

them, she arrogates to herself a privilege—the free enjoyment of those benefits—while depending on the renunciation of that privilege by the others.

Now these remarks, I readily accept, are not in themselves especially illuminating. The claim is not that in all unfair actions, the agent preferentially favors herself;³⁵ but rather that in this case, the unfairness consists in objectionably preferential treatment. And there is obviously no question of entertaining any analytic ambitions for the latter phrase—the sense in which the preferential treatment is objectionable is that it is unfair. So far, then, my thoroughgoing opponent, who finds nothing unfair in the refusal to pay for nonvoluntary benefits, has been given no reason to disagree. The contentious issue is this: In just what feature of the Fare-Evader's conduct does her objectionably preferential treatment of herself lie?

One feature of obvious relevance is her dependence for the benefits she receives on others' willingness to pay, without being willing to do so herself. Once more, there is no need for my opponent to deny this. What he must deny, though, is that this feature is by itself sufficient to make a person's treatment of herself objectionably preferential. For such a dependence is exhibited by agents (such as the Recalcitrant Fisherman) who refuse to pay for nonvoluntary benefits, as well as by those who refuse to pay for benefits they have sought out. However, there is surely a sense in which any case exhibiting a dependence of this sort is one where the agent gives herself preferential treatment: she makes a special case of herself, allowing herself not to pay for goods that she either does or ought to realize are worth paying for, and which she only receives because others are moved by the same realization to pay. The issue between me and my opponent is whether preferential treatment in *this* sense is objectionable. As far as I can see, making a special case of oneself in this way does intuitively amount to unfairness. But my thoroughgoing opponent disagrees. For him, we have only identified a sense in which the Fare-Evader's treatment of herself is *objectionably* preferential when we have invoked the fact that she deliberately *takes* the benefits for the existence of which she depends upon others' willingness to pay.

The question I wish to press against the thoroughgoing opponent is

35. In no straightforward sense is this true of a judge who becomes bored with the litigants before him and tosses a coin, for example.

this. *How* does one's deliberately taking, rather than merely receiving, a benefit contribute to making one's treatment of oneself objectionably preferential? I shall consider the various answers open to him, and argue that none is satisfactory.

The quickest sort of answer will be that taking the benefits involves *theft*, or the violation of *rights*. But neither of these suggestions will get us very far, even supposing a close connection between them and considerations of unfairness. Clearly, the unauthorized taking of benefits is not always morally objectionable: consider a bird-watcher who refuses to buy from a fruit-grower outside whose orchard he has come to camp. We need to supply the grounds for regarding the taking of benefits by the Fare-Evader as objectionable, and thus a candidate to be described as theft or the violation of rights, but other cases of the unauthorized taking of benefits as unobjectionable. These are not moral premises, but moral conclusions.

A more promising attempt to show how the Fare-Evader's taking the benefit contributes to her unfairness appeals to differences between the content of the intention of the taker of a benefit and that of a receiver. That there are such differences is undeniable: the Fare-Evader intends to take the benefit, whereas the Recalcitrant Fisherman does not—it is forced upon him against his will. And this is to say that the Fare-Evader allows her practical plans to be governed by the actions of compliers in a way that the Recalcitrant Fisherman's are not. Had the other commuters acted differently, and the bus-riding benefits been unavailable, the Fare-Evader would have had to modify her actions; had the other fishermen acted differently, not conferring the benefits on the Recalcitrant Fisherman, he sincerely tells us that he would simply have carried on as before.³⁶ So while there may be a sense (the one mentioned above) in which both of them *depend* for the receipt of benefits on others' compliance, there is also a sense in which only the Fare-Evader *relies* on it. And this provides a sense in which the Fare-Evader *takes advantage of*, and *exploits*, the others' benefit-producing compliance, but the Recalcitrant Fisherman does not.³⁷

36. For a full account of the connection between intentions and plans toward which I am gesturing here, see Michael E. Bratman, *Intention, Plans, and Practical Reason* (Cambridge: Harvard University Press, 1987).

37. I am grateful to the Editors of *Philosophy & Public Affairs* for this formulation of the point.

There may be such a sense, but it is not yet one in which taking advantage of and exploiting other people is unfair. For it applies equally to the actions of any fare-payer. Paying commuters share the Fare-Evader's intention of securing the benefits of the public transport scheme—their plans are accordingly governed by the actions of other compliers, in the same way—and yet they do not act unfairly. Now this remark may seem to be merely beside the point: no one is saying that the taking of benefits contributes to making an action unfair, irrespective of whether the benefits are paid for. What it shows, however, is that if there is a morally significant sense in which the Fare-Evader takes advantage of, or exploits, the others' compliance, it is not provided simply by her intending to take the benefits it produces. To secure a morally significant sense in which she does so, my opponent clearly must invoke not only the intention to secure the benefits, but nonpayment as well.

But how can he do so? It will not help to describe the Fare-Evader's intention more carefully as one of *taking the benefit without paying*. This suggestion avoids the previous problem—the sense in which someone with this intention exploits the compliers who produce the benefit does now seem morally significant. Its problem, though, is the converse: this intention is one the Fare-Evader need not possess. Perhaps she has got onto the bus intending to pay, because she expects to be asked to do so by a conductor; when none appears, and once she is already receiving the benefit, she decides to see whether she can get away without paying. Here, the Fare-Evasion is not premeditated, but it surely exhibits the same sort of unfairness as the premeditated variety. At no time does the unpremeditated Fare-Evader intend to take the benefits without paying: when she first forms the intention of not paying, they have already been taken.

Thus it is difficult to see how appealing to the content of the Fare-Evader's intention helps to advance my opponent's case, according to which her taking the benefits is essential to the unfairness of her conduct. Either he underspecifies the content of that intention—simply as *taking the benefits*—with the result that fare-payers share it, or he overspecifies its content—as *taking the benefits without paying*—so that not all Fare-Evaders possess it. Of course, there is an intention that all Fare-Evaders share, and which I am happy to allow yields a morally significant sense of exploitation: the *de re* intention not to pay for benefits that have only been produced through others' willingness to do so. But *this*

intention is displayed equally in the Recalcitrant Fisherman's refusal to pay for the benefits nonvoluntarily conferred upon him. In *this* sense, he too exploits, or takes advantage of, the willingness of others to pay for the benefits he receives—he deliberately refuses to pay for them—and this, I am maintaining, is the morally significant sense.

When I say this, I am not denying that the Fare-Evader's intentions are different from those of both fare-payers and Recalcitrant Fishermen. The difference can easily be stated: she possesses, at the same or different times, both the intention to take the benefits, and the intention to refrain from paying. What I have been arguing, though, is that it is hard to see how the first intention makes a further contribution toward her *unfairness*, when it is added to her unwillingness to pay for benefits for the receipt of which she depends on others' willingness to pay.

But perhaps there is a more attractive line for my opponent to take. Perhaps it is not their intentions concerning their own actions that secure him the morally significant difference between the Fare-Evader and the Recalcitrant Fisherman, but their preferences concerning others'. The Recalcitrant Fisherman, as initially described, was prepared to extend the privilege of nonpayment to everyone else. He would actually *prefer* it if no one contributed; and therefore (it might be maintained) does not arrogate any privileges peculiarly to himself.

The difficulty now, however, is to see what prevents the Fare-Evader from possessing a similar willingness to extend her privilege of nonpayment to other bus-users. It might appear that since her conduct plainly exhibits a preference for acquiring the benefit, the profession of such a willingness could not be sincere. But what stands in the way of sincerity here? Can she not prefer it even more if everyone shared her disposition not to pay, and there was consequently no benefit to be had? Not if she is *rational*, to be sure, for this would leave everyone, including her, worse off. It does appear that rational preferences about the conduct of others, consistent with the Fare-Evader's own conduct, would commit her to an arrogation of privileges of nonpayment to herself. But once more, it will not help the opponent to cite *this* feature of the case of Fare-Evasion as contributing toward its morally objectionable nature. For again, it is exemplified by the Recalcitrant Fisherman. His preference, sincere though it may be, is irrational, given the fundamental benefit conferred on him by the preservation of his livelihood. It is tempting to protest that although the Fisherman is *given* preferential treatment by

the others, he cannot be accused of *giving himself* preferential treatment, since he does not give himself anything. But the response should be clear. It is not by taking the benefits that he is giving himself preferential treatment, but by refusing to pay for them.

It is true, of course, that as initially described, the Fare-Evader resorts to *deception* and the Recalcitrant Fisherman does not: it is essential to the success of her plan that she does not announce her intention of nonpayment in the way that he does. And this is surely one morally objectionable feature of her conduct. But it is not the one my opponent is after; for a more brazen Fare-Evader who makes no secret of her intention to free-ride on the others still exhibits the unfairness he is seeking to explain.

So it is hard to see how the deliberate taking of benefits helps to characterize the Fare-Evader's objectionably preferential treatment of herself, in which the unfairness of her conduct consists. And if, in such paradigmatic cases of free riding, even the *taking* of benefits does not contribute to characterizing their unfairness, then it is hard to see how the voluntary acceptance of benefits can ever do so. The feature of primary moral significance, it seems, is simply the Fare-Evader's dependence for the benefits she receives on others' willingness to pay for them, without being willing to do so herself. This can be displayed equally in the refusal to pay for nonvoluntary benefits. In the absence of any explanation of why this feature should not support the unfairness of the latter conduct if it does the former, we should accept that it, too, is unfair.³⁸

Do not misconstrue the method of argument I have been using to this end. My question has been, "What is it about the taking of benefits that helps it to contribute to the unfairness of Fare-Evasion?" I have been considering in turn the features of it that might be cited in reply, and showing that in each case, the feature is also present in cases where either there is no unfairness or benefits are conferred nonvoluntarily. But I am not arguing that if in one case possessing a certain feature, there is no unfairness, then that feature cannot contribute to the unfairness of Fare-Evasion. (This would be fallacious, relying on the false assumption that if feature A of a first situation and feature B of a second

38. This formulation of the argument should be acceptable even to moral "particularists" like Jonathan Dancy, *Moral Reasons* (Oxford: Blackwell, 1993), for whom the fact that a certain feature of a situation is a reason for its unfairness does not entail that that feature is a reason for the unfairness of any other situation in which it is present.

are without moral significance then they cannot be morally significant in conjunction with each other.)³⁹ Rather, I am arguing that if so, then the feature is not sufficient on its own for unfairness, and hence that its purported contribution to unfairness is not obvious. We require a further explanation of how the feature's combining with others does contribute to unfairness; but the further explanations have been found wanting, in the same way.

The Principle of Fairness ought, therefore, to be endorsed in a version in which it is extended to encompass nonvoluntary benefits. But why (finally) in the detailed way advocated in Section III above? The foregoing characterization of the free rider's unfairness equips me with the only materials I need to answer this. The core of the principle is the claim I have just been discussing: if a person receives a net benefit from a scheme, then her unpreparedness to meet its requirements, when she depends for the benefits she receives on others' meeting them, is unfair. The main elaborations on this core are the introduction of the notion that it is the *practice of participation* in the scheme that is to represent a net benefit for her, and the addition of two further conditions. The support for these elaborations can now be stated, by recalling the earlier discussion.

To show the need for the notion of "the practice of participation," I began by considering the requirement that in order for nonpayment for benefits to be unfair, they must be worth their cost. Why should *this* be required? Because the refusal to pay, if such a requirement is not met, displays no preferential attitude toward my own interests. Either I am being treated worse than others, or everyone would be better off without the scheme's costs and benefits: in neither case does the refusal to cooperate display an arrogation of privileges to myself. I then observed that there is a sense in which a luxurious benefit compulsorily supplied to a poor person might be worth its cost, but that this sense appeared irrelevant. And a simple explanation can surely be offered of *why* this sense is irrelevant in the same terms: if the poor person would be *worse off* getting the benefits and paying, but a rich person would not, then the poor person's refusing to pay displays no preferential attitude toward his own interests.

39. Consider the features of inflicting pain, and doing something because of the enjoyment one derives from it. The fallacy is identified in Shelley Kagan, "The Additive Fallacy," *Ethics* 99 (1988): 5–31.

This led me to distinguish a second sense in which the benefits conferred by a scheme may be worth their cost: I can be better off getting the benefits and paying than I am doing neither. But three refinements had to be made to this requirement. First, the benefits that are actually attributable to the scheme must be circumscribed, and merely fortuitous consequences of its operation, such as the discovery of treasure, ruled out. Secondly, we must take into account those costs that the scheme imposes and that are relevant to the justification of nonpayment, such as those imposed in making the fisherman's water unsafe. And thirdly, benefits attributable to the practice of *non*participation in the scheme must be excluded from consideration.

In each case, the refinement I proposed is supported by my simple characterization of the kind of unfairness exhibited by the free rider: her objectionably preferential treatment of herself, in allowing herself not to pay for goods that she either does or ought to realize are worth paying for, and that she only receives because others are moved by the same realization to pay. The relevant benefits must be those at which participants in the scheme are aiming, since where this is not the case (as with the treasure), the beneficiary is not dependent on others' willingness to pay for *them*. The same point obviously explains the exclusion of benefits attributable to the practice of nonpayment—indeed, here, a beneficiary is not merely not dependent on compliance with the scheme's requirements for these benefits; she is dependent on noncompliance. Why should we identify the relevant costs with those for which the participants are collectively morally answerable—for which participants in the scheme are required to offer either a morally good reason or compensation? Approach the question this way. Suppose I am accused of free riding on a scheme, and adduce certain costs that result from its operation, and that outweigh its benefits to me, in my defense. For which sorts of costs will that defense meet the accusation? On the proposed characterization of the free rider's unfairness, the accusation is that I give myself objectionably preferential treatment. Now suppose the costs I cite are ones for the imposition of which I can require participants in the scheme to produce a justification. If I can require this, what they offer as that justification must be that the costs are unavoidable if the scheme's benefits are to be produced. And if so, then when those costs outweigh the benefits, my requirement has not been met: as before, I am either being treated worse than others, or everyone would be

better off without the scheme; either way, the accusation of objectionably preferential treatment fails. But if the costs I cite are not associated with such a requirement, then I am simply not speaking to the accusation. I have been challenged to justify my failure to pay for certain benefits, and have responded by citing a separate set of costs. The accusation of objectionably preferential treatment stands unaddressed. So given the proposed characterization of the free rider's unfairness, the costs that are relevant to the justification of nonpayment are those associated with this requirement.

Beyond this, my Principle of Fairness adds two further conditions. It is to be the case neither that the fair generalization of the scheme's requirements, in the manner discussed earlier, is to the detriment of practically everyone; nor that I am raising a legitimate moral objection to the scheme. How do these conditions contribute to the case for regarding nonparticipation as unfair? They do so for reasons of the same general form. The inclusion of the first is justified if the fact that the fair generalization of the scheme would be detrimental to practically everyone suffices (even if I'm actually receiving a net benefit) to make it fair for me to refuse to comply. That is, if I can fairly refuse to comply whenever our all regarding as obligatory all requirements of a certain sort—those that must in fairness be so regarded if the scheme's requirements are themselves regarded as obligatory—would be detrimental to practically everyone. Why should we think so? Because if *practically everyone* will be better off if we all ignore demands of this sort, then my doing so involves no arrogation of preferential advantages. I am not objectionably favoring myself if I am only doing something everyone's doing which would be to practically everyone's advantage. Why should the relevant sort be picked out in this particular way (by reference to what fairness requires us to regard as comparable)? Because if I tried to argue this on the basis of some other, self-favoring standard of generalization, then I *would* be arrogating such advantages. And why adopt the specific reading of "practically everyone" given earlier (as meaning either everyone, or enough others to make it unreasonable for the minority to resist the majority-beneficial outcome)? Because if a minority who would be worse off under the fair generalization of a scheme can fairly resist the majority-beneficial outcome, then there *is* unfairness in my endorsing this course of conduct for everyone.

Finally, the connection with considerations of preferential treatment also supports the inclusion of condition (iii), concerning the raising of legitimate moral objections. If there are such objections, then I should not be preferentially favoring myself in refusing to support it on such grounds, since everyone else ought to be doing the same.

Thus all the clauses of the Principle of Fairness I am proposing can be seen to flow naturally from a simple characterization of the kind of unfairness displayed by the Fare-Evader. The question with which my discussion began—"If the free rider harms no one, what is it about her conduct that makes it unfair?"—has now been fully answered, and both of the initial tasks completed. The features of her conduct that make it unfair are those identified by my principle; and those features *make* it unfair by convicting her of the sort of objectionably preferential treatment often characteristic of unfairness.

V. PRINCIPLES AND JUSTIFICATION

I close with some remarks on two subjects: the justification that has just been offered for my principle, and the principle's own justificational ambitions.

The argument of the preceding section should not be misconstrued as a foundational one. It seeks to support the various clauses of the principle by invoking a general characterization of the Fare-Evader's unfairness in terms of the "arrogation of preferential advantages." But there is obviously no question of using a phrase this imprecise to *deduce* anything about the exact circumstances under which a person's conduct is unfair. Giving a precise interpretation to the phrase is hardly something that can be done in advance of thinking about particular cases of unfairness—and arriving at the sorts of intuitive judgments which are themselves summarized by the principle. The claim of the previous section is that the formulation of the principle is guided by a coherent and recognizable conception of fairness—a conception that the principle itself helps to clarify.

Finally, let me say something about the status of this principle. Had it purported to supply a set of "descriptive" conditions upon which unfairness supervenes, my proposal would have been hopelessly inadequate. After all, its second condition itself makes reference to fairness,

and the third to moral legitimacy more generally. What has been presented might therefore be thought question-begging, or at least unhelpful. However, the intention has not been to produce a description that would furnish someone antecedently incapable of moral discriminations with an understanding of what the free rider's unfairness consists in. My aim has only been to state what people concerned about fairness are reacting to, in characterizing free riding as unfair. There are different species of unfairness: my account of one of them mentions others. If there is a correct reductive account of those further species of unfairness—and the remarks concerning “objectionably preferential treatment” surely do not supply it—it may simply be inserted into my principle.

Nor, moreover, need any dubious justificational ambitions be embodied in my advocacy of this principle. It should not be thought that I have committed myself to a picture of the justification of moral commitments or moral judgments as the deduction of particular judgments from universal principles of this kind. An alternative is to say that the *justification* of anyone concerned to avoid free riding, in an uncomplicated case, need only be that he benefits from a scheme that produces benefits by making requirements of beneficiaries. There are of course many considerations that could countervail against this one, relating to circumstances in which nonpayment would be fair: the principle summarizes them. Indeed, I would prefer this latter way of conceiving of the justification of moral judgments, commitments and action. But defending it is another matter:⁴⁰ the point here is only that preserving the interest of the principle is compatible with an uncommitted stance on the structure of ethical justification. It claims to summarize the conditions under which one is justified in thinking that nonpayment is unfair, but need not itself constitute the justification.

APPENDIX: PUBLIC GOODS

This brief review of the literature on public goods has two aims. The first is to show that the only thing definitions of public goods seem to have

40. For a fuller account of this conception of moral justification, see Garrett Cullity, “International Aid and the Scope of Kindness,” *Ethics* 105 (1994): 99–127, and idem., “Aretaic Cognitivism,” *American Philosophical Quarterly* (forthcoming).

in common is their involving some subset of the seven features I earlier called “Jointness in Supply,” “Nonexcludability,” “Jointness in Consumption,” “Nonrivalness,” “Compulsoriness,” “Equality,” and “Indivisibility” (see pp. 3–4 above). The second is to give the provenance of these labels.

The *locus classicus* of the contemporary economic public goods literature is Paul A. Samuelson, “The Pure Theory of Public Expenditure,” *The Review of Economics and Statistics* 36 (1954): 387–89. (For antecedents, see J. G. Head, “Public Goods and Public Policy,” *Public Finance* 17 [1962]: 197–98.) He emphasizes the third feature and infers from it the seventh; but the inference is dubious if those to whom a good with the third feature is available can have different capacities for consuming it. If the seventh feature is made definitive of publicity, however—as Samuelson, “Diagrammatic Exposition of a Theory of Public Expenditure,” *The Review of Economics and Statistics* 37 (1955): 350–56, suggests—the third can be inferred from it. Head, “Public Goods,” pp. 198–206; Geoffrey Brennan, “Economics,” in Robert E. Goodin and Philip Pettit, eds., *A Companion to Contemporary Political Philosophy* (Oxford: Blackwell, 1993), pp. 144–45; and John G. Cullis and Philip R. Jones, eds., *Public Finance and Public Choice: Analytical Perspectives* (London: McGraw-Hill, 1992), pp. 60–63, follow this pattern—although their claim that the seventh feature also entails the second is questionable. Ledyard, “Incentive Compatibility”; and Pearce, ed., *Macmillan Dictionary*, p. 352, make the third feature definitive of publicity; Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986), pp. 198–99, the second; Lawrence C. Becker, “The Free-Rider Problem,” in Harlan B. Miller and William H. Williams, eds., *The Limits of Utilitarianism* (Minneapolis: University of Minnesota, 1982), p. 217, the first; Rutherford, *Dictionary*, p. 375, the first three; Miller and Sartorius, “Population Policy,” p. 151, the first two; David Schmidtz, *The Limits of Government* (Boulder: Westview, 1991), p. 55, and Cowen, ed., *Public Goods*, p. 3, the second and third; and J. Richard Aaronson and Attiat F. Ott, “The Growth of the Public Sector,” in David Greenaway, et al., eds., *Companion to Contemporary Economic Thought* (London: Routledge, 1991), p. 529, the second, third, and sixth. Other writers prefer to use these features to define pure publicity, and make publicity a matter of the degree to which this is approximated. For J. M. Buchanan, *The Demand and Supply of Public*

Goods (Chicago: Rand McNally, 1968), p. 49, the second feature is definitive of pure publicity; for Francis M. Bator, "The Anatomy of Market Failure," *Quarterly Journal of Economics* 72 (1958): 369, the third; for C. V. Brown and P. M. Jackson, *Public Sector Economics*, 4th ed. (Oxford: Blackwell, 1990), p. 34, the seventh; for Pearce, ed., *Macmillan Dictionary*, p. 352, the second and third; and for Arneson, "Principle of Fairness," the second, third, and sixth.

There is as little unanimity in the names given to the seven features as there is concerning the term "public goods" itself. I take the first four from others; the last three are my own coinages.

For "jointness in supply," see Miller and Sartorius, "Population Policy," p. 151, and Hampton, "Free-Rider Problems," 247. Schmitz, *Limits*, p. 55, calls the first feature "nonexclusiveness," and Klosko, "Principle of Fairness," p. 354, "nonexcludability." (Notice the difference between it and the following entailment of the seventh feature, which Head, "Public Goods," p. 201, calls "jointness": "once produced, any given unit of the good can be made equally available to all.") For "nonexcludability," used of the second feature, see Miller and Sartorius, "Population Policy," p. 151; Cowen, *Public Goods*, p. 3; and Arneson, "Principle of Fairness," pp. 618–19, who also calls goods with this feature "collective." The label "jointness in consumption" might more appropriately be applied to the fifth, sixth, or seventh feature. For its use in relation to the third, though, see Brennan, "Economics," p. 145, who also calls this "nonrivalness," not distinguishing it from the fourth feature. Samuelson, "Pure Theory," p. 387, calls this "collective consumption"; Arneson, "Principle of Fairness," p. 618, Anthony de Jasay, *Social Contract, Free Ride* (Oxford: Clarendon Press, 1989), p. 157, and Michael Taylor, *The Possibility of Cooperation* (New York: Cambridge University Press, 1987), p. 7, simply "jointness." For "nonrivalness," used of the fourth feature, see Taylor, *Possibility*, p. 7. It is more usually attached to the third, though: see e.g. Cowen, ed., *Public Goods*, p. 4, Rutherford, *Dictionary*, p. 375, and Pearce, ed., *Macmillan Dictionary*, p. 352. "Compulsoriness" is my label for what Brown and Jackson, *Public Sector*, p. 35, call "non-rejectability," and Simmons, "Principle of Fair Play," p. 327, calls "openness." "Equality" is my label for what Arneson, "Principle of Fairness," pp. 618–19, makes the third feature of "pure public goods," and Rawls, *Theory of Justice*, p. 266, calls "indivisibility." I am using "indivisibility" for what Samuelson, "Diagrammatic Exposition," p. 350, calls "public consumption."