**Utilitarianism and Fairness**

forthcoming in Ben Eggleston and Dale Miller (eds.) *Cambridge Companion to Utilitarianism* (2014)

B.W.hooker@Reading.ac.uk

Those who are not skeptics about fairness are almost always in favour of it. And one of the most prominent objections to utilitarianism is that it ignores, and its requirements can conflict with, fairness. In order to discuss such ideas, we need to distinguish between different kinds of utilitarianism and between different ideas about fairness. In the next two sections, I make these distinctions. In the subsequent two sections, I discuss the possibility of conflict between different kinds of fairness and different kinds of utilitarianism.

**Section 1: Different kinds of utilitarianism**

When early utilitarians referred to “the greatest good of the greatest number,” they might have had in mind not only the maximization of aggregate welfare but also the spreading of benefits to as many individuals as possible. In cases where what would produce the greatest aggregate welfare would not be what would benefit as many individuals as possible, which is more important? Whatever answer early utilitarians would have given, at least since 1960 the term ‘utilitarianism’ has usually referred to a moral and political philosophy that evaluates either acts or rules or both purely in terms of the aggregate welfare. Utilitarian calculation of aggregate welfare is impartial in that benefits or harms to any one individual are counted exactly the same as the same size benefits or harms to any other individual.

The kind of utilitarianism most often discussed in textbooks is the kind that conflicts most sharply with non-utilitarian theories. This kind of utilitarianism is maximizing act-utilitarianism. According to maximizing act-utilitarianism, (a) an act is morally *required* if, only if, and because it would produce greater aggregate welfare than any alternative act, (b) an act is morally *permissible* if, only if, and because no alternative act would produce greater aggregate welfare, and (c) an act if morally *wrong* if, only if, and because at least one other alternative act would produce greater aggregate welfare.

In the rest of this chapter, I will regularly refer just to a theory’s account of what is moral required, or of what is morally permitted, or of what is moral wrong, but not all three. This is just to save words. Each theory I will discuss will have an account of all three.

Another kind of act-utilitarianism is satisficing act-utilitarianism, which is the view that an act is morally permissible if, only if, and because the act produces enough aggregate welfare, the presumption being that enough aggregate welfare is something short of the maximum available.[[1]](#footnote--1) Satisficing act-utilitarianism is like maximizing act-utilitarianism in holding on to the distinctions between the morally permissible, the morally required, and the morally wrong. But satisficing act-utilitarianism draws the distinctions in different places than maximizing act-utilitarianism does. For maximizing act-utilitarianism, the line between permissible and wrong is drawn right at the top of the scale of good outcomes—only acts which produce at least as much utility as any other act are morally permissible. For satisficing act-utilitarianism, in order to be permissible, acts have to produce enough utility but don’t have to produce the maximimum available.

One of the great attractions of satisficing act-utilitarianism in comparison with maximizing act-utilitarianism is that satisficing act-utilitarianism leaves the agent with a larger range of morally permissible alternatives. In every situation, maximizing act-utilitarianism shrinks the set of morally permissible actions down to the single act that maximizes welfare, unless there two or more equally maximal acts, in which case maximizing act-utilitarianism requires the agent to choose one of these. Shrinking the set of morally permissible actions so far is implausibly restrictive, leaving the agent with vanishingly little moral freedom to choose between morally permissible acts.[[2]](#footnote-0) In contrast, depending on where satisficing act-utilitarianism draws the threshold of “good enough,” there will be many more than one or two possible acts that pass this threshold and thus qualify as morally permissible according to satisficing act-utilitarinism.

The other great attraction of satisficing act-utilitarianism is that it typically requires less self-sacrifice of the agent than does maximizing act-utilitarianism. Maximizing act-utilitarianism requires the agent to keep sacrificing for the sake of others up to the point where further sacrifices will harm the agent more than they will benefit others. Satisficing act-utilitarianism requires merely that the agent make whatever sacrifices are needed to produce enough aggregate welfare, but this will typically be somewhat less than what would be required to maximize aggregate welfare.

While these two attractions of satisficing act-utilitarianism are very powerful, there are some devastating objections to satisficing act-utilitarianism. The most obvious difficulty with satisficing act-utilitarianism is about how it can draw a *non-arbitrary* line between enough aggregate welfare and not enough aggregate welfare. Even more compelling is the objection Tim Mulgan advanced against satisficing act-utilitarianism that it permits the agent to choose what benefits others less than the maximum possible even when choosing what would benefit them the most would be no more costly or difficult for the agent.[[3]](#footnote-1)

Another kind of act-utilitarianism is scalar act-utilitarianism.[[4]](#footnote-2) This view jettisons the distinctions between morally required, morally permissible, and morally wrong in favor of an undifferentiated scale from best possible act to worst possible act. Thus, according to scalar act-utilitarianism, no acts are morally required, morally permissible, or morally wrong. Scalar act-utilitarianism holds instead that there is the act that would produce the greatest aggregate welfare, and there is the act that would produce the next greatest aggregate welfare, and there is the act that would produce the next greatest aggregate welfare, and so on down the scale to the act that would result in the least aggregate welfare.

The advantage scalar act-utilitarianism is supposed to have over maximizing act-utilitarianism is that scalar act-utilitarianism does not have the burden of defending the extremely restrictive and demanding standard of morally permissible action that maximizing act-utilitarianism avows. One advantage scalar act-utilitarianism has over satisficing act-utilitarianism is that scalar act-utilitarianism does not have the burden of specifying what counts as *enough* aggregate welfare. Another advantage of scalar act-utilitarianism over satisficing act-utilitarianism is that scalar act-utilitarianism does not have to admit that producing less than the maximum benefits to others is morally permissible when producing the maximum would involve no greater sacrifices for the agent.

But scalar act-utilitarianism can take such positions only because it has abandoned altogether the concepts of moral permissibility and moral wrongness and thus the distinction morally permissible acts and morally wrong acts. This distinction is absolutely central to moral thought. Much moral thought is focused on what moral duty requires of us, as opposed to what is permitted but not required, for example because it goes beyond the call of moral duty. And many moral reactive attitudes pivot on the distinction between wrongness and permissibility. For example, an act cannot be morally blameworthy unless it was morally wrong; and feelings of guilt, resentment, or indignation are out of place if what was done was morally permissible.

From now on, the only variety of act-utilitarianism I will discuss is maximizing act-utilitarianism. Maximizing act-utilitarianism has been far more prominent than satisficing act-utilitarianism and scalar act-utilitarianism. Furthermore, the objections to satisficing act-utilitarianism and scalar act-utilitarianism seem to me completely compelling.

Act-utilitarianism can be formulated in terms of actual results or in terms of probabilities. Act-utilitarianism formulated in terms of actual results holds that these results determine what was morally required, permissible, or wrong. Act-utilitarianism formulated in terms of probabilities starts by listing the possible benefits and possible harms of an act. Then it multiplies those possible benefits and harms times the probabilities of their occurring if the act is done. Then it sums these numbers to reach the “expected utility” of the act.

Here is a standard example. Suppose a doctor could give her patient a pill that has a 99.9% chance of curing what would otherwise be a fatal cancer and a .1% chance of killing the patient immediately. Suppose the doctor then gives the patient the pill and the patient dies. The actual utility of giving the patient the pill was very low, indeed very negative. But the expected utility of giving the patient the pill was very high.

Of course we would regret the doctor’s giving this patient this pill. But we certainly wouldn’t blame the doctor for doing what she sincerely and reasonably believed had a 99.9% chance of curing the patient. We might even say that it would be *unfair* to blame the doctor for being so terribly unlucky. In fact, we blame people for taking terrible risks with other people’s welfare, even when luck prevailed and the probable harms did not in fact occur.

Such thoughts militate in favor of casting any discussion of relations between utilitarianism and fairness in terms of expected utility instead of actual utility. This is what I shall do.

Now if utilitarianism is understood as a family of theories each of which evaluates either acts or rules or both purely in terms of the aggregate welfare, how should we classify the member of this family that evaluates *both acts and rules* purely in terms of aggregate welfare? In some circles, this theory is called *global utilitarianism*.[[5]](#footnote-3)

Suppose the rule whose acceptance would maximize aggregate welfare could, in at least some situations, require of the agent an act that would not maximize aggregate welfare. Global utilitarianism would endorse accepting this rule, since by hypothesis its acceptance is welfare maximizing. At the very same time, global utilitarianism would condemn acts of compliance with the rule in these situations, since by hypothesis those acts are not welfare maximizing.

Thus, global utilitarianism is a kind of act-utilitarianism. Now, many prominent act-utilitarians, e.g., Henry Sidgwick and J. J. C. Smart, have been global utilitarians.[[6]](#footnote-4) Could one be an act-utilitarian without being a global utilitarian? Yes, one could be an act-utilitarian about acts and yet assess rules not in terms of the utility of their acceptance but in terms of the frequency with which these rules would result in right (utility-maximizing) acts. But such possibilities need not bother us here. For the purposes of this chapter, what matters about global utilitarianism is its act-utilitarian assessment of acts.

Unlike all forms of act-utilitarianism, including global utilitarianism, rule-utilitarianism assesses acts in terms of rules. According to rule-utilitarianism, the rules with the greatest expected utility are justified, and acts are justified if and because they are allowed by these rules.

If poorly formulated, rule-utilitarianism runs into fatal objections. On some formulations, rule-utilitarianism does not even manage to disagree with maximizing act-utilitarianism about which actions are morally right. However, this chapter is not about how to formulate rule-utilitarianism, nor about the multitude of possible objections to rule-utilitarianism.[[7]](#footnote-5) In the present chapter, without explaining why, I will simply presuppose that the best formulation of rule-utilitarianism is that an act is morally permissible if, only if, and because the act is allowed by the code of rules whose internalization by the overwhelming majority of everyone in future generations would maximize expected aggregate welfare. And, to save words, I will often use the term “optimific rules” to refer to rules whose internalization by the overwhelming majority of everyone in future generations would maximize expected aggregate welfare.

**Section 2: Different kinds of fairness**

Just as we need to distinguish between different kinds of utilitarianism, we need to distinguish between different kinds of fairness.

Rules and principles make distinctions. Formal fairness consists in the equal and impartial application of these distinctions. For example, if there is a rule that people who earn over 100,000 have to pay 40% in tax, then formal fairness is violated if the senator’s husband is allowed to pay a lower rate of tax though he earns over 100,000. If there is a rule that the first one in line gets served first, then formal fairness is violated if, though the first in line is normally served first, this isn’t true when the first in line has an unpopular accent.

The equal and impartial application of the same rules to everyone—formal fairness—is not exhaustive of fairness, since rules applied equally and impartially to everyone can be unfair in their content. Whether or not rules are applied equally and impartially, they can make distinctions that fairness forbids. To illustrate, imagine a set of rules distributing advantages or disadvantages purely on the basis of eye color or gender. The equal and impartial application of such rules would be *formally fair* but *substantively unfair*. In order for rules to have fair content, the distinctions between people made by these rules must be ones that fairness allows to be made or even demands to be made. Distinguishing between people on the basis of eye color or gender is not allowed by fairness.

It is much easier to show that we must distinguish between formal and substantive fairness than to say what determines substantive fairness. There are a number of rival views. I catalogue the most prominent in the following section.

**Section 3: Different views of substantive fairness**

A social practice is partly constituted by rules, which then frame expectations, intentions, reactions to noncompliance, etc.[[8]](#footnote-6) Social practices have many different purposes, of course. Nevertheless, another common idea is that a social practice’s rules are fair depending on whether they make distinctions that serve the point of the social practice.

Formulated broadly as a claim about *every* social practice’s rules, that idea is implausible. For some social practices have indefensible purposes. A rule could not be rendered fair simply by the fact that it makes distinctions that serve the point of an indefensible social practice. Imagine a social practice the purpose of which is to make people feel superior to others or inferior to others on the basis of the social standing of their parents. The rules of this social practice might dictate that anyone whose parents did not have high social standing should bow and defer to anyone whose parents did have high social standing. This rule would serve the purpose of the social practice in question, but would not be making a distinction that fairness allows. Fairness does not allow a distinction made on the basis of the social standing of people’s parents.

So we cannot accept the broad idea *any* social practice’s rules are fair depending on whether they make distinctions that serve the point of the social practice. We should thus consider a narrower idea focused on social practices with defensible and thus permissible purposes. This is the idea that any distinctions that serve permissible purposes of social practices are permitted by fairness.

I do not mean to suggest that social practices have permissible purposes only when those purposes serve fairness. Many social practices have the purpose of efficiently producing goods that are specifiable independently of the social practice and do not include fairness. Such goods might be knowledge or health or enjoyment. Many social practices have the purpose of paying homage to independently specifiable goods. An example is the social practice of congratulating teachers on the success of their former students. And many social practices have the purpose of specifying a kind of excellence and the conditions under which this excellence can be achieved. Athletic and intellectual games and artistic genres of many sorts provide examples.

Since social practices are permitted to have so many different purposes and many of these overlap or complement one another, consider the idea that society is a network of social practices. Must society achieve certain things in order for its social practices (and their constituent rules) to be fair? In the remainder of this section, I will outline a number of the possible alternative answers to this question. In later sections, I will say how conceptions of fairness can conflict with act-utilitarianism or with rule-utilitarianism.

One idea is the social practices of a society are fair if and only if *the social practices have maximum expected utility*. The line of thought might be that utilitarian assessment of social practices and rules has, at the foundational level, equal concern for the welfare of each. Since the foundation seems so eminently impartial, is not this foundation fair? If this is a fair foundation, then would not the social practices and the acts those social practices license inherit fairness from the foundation?

A second and very different idea is that the social practices of a society are fair if and only if the social practices require each to contribute equally to public goods as long as others are contributing. Whereas the previous idea suggested that fair practices are whichever ones maximize expected utility, the second idea ties fairness to *reciprocity*. The complaint that existing social arrangements allow “free loaders” and “tax cheats” to piggy back on the hard work of others testifies to the underlying appeal of this conception of fairness.

A third and again very different idea is that the social practices of a society are fair if and only if they are geared to make sure that people get what they *need*, at least where the society has enough resources to provide this. It is often said to be manifestly unfair that some have more than they need while others have less than they need.

A fourth idea is that social practices are fair if and only if they *leave the worst off as well off as the worst off under any other arrangements would be left*. This idea was made popular by John Rawls.[[9]](#footnote-7) His “difference principle” held that inequalities are fair only if they leave the worst off better off than the worst off without such inequalities would be left. Rawls’s difference principle implies the requirement to “*maximin*” (for “maximize the minimum position”).

Rawls’s “maximize the minimum” was justified, Rawls argued, by the manifest fairness of a rational choice situation that Rawls took over from John Harsanyi and then modified. Suppose that, instead of trying to select rules from a point of view in which there is equal concern for everyone and full information about the likely benefits and harms of different possible rules, we select rules for society from an “original position” in which we care about only ourselves but are behind a “veil of ignorance” which hides from us all specific information about ourselves. Behind this veil, we don’t have any idea whether we are talented, energetic, healthy, female, religious, etc. Since behind the veil we have no information that could bias our selection of rules, Rawls proposes the rules we select behind the veil would be selected fairly and impartially.

Harsanyi argued that the rational choice behind the veil of ignorance would be to choose whatever rules would maximize utility. If one had an equal chance of being anyone once the veil went up, then the way to maximize one’s own expected utility behind the veil would be to pick the rules with the greatest expected average utility, everyone’s utility being counted equally and impartially. Rawls argued against Harsanyi, however, that behind the veil of ignorance one would be rational to be risk averse and thus focus on the position of the worst off instead of the average position.

That argument of Rawls’s has widely been thought to be unpersuasive. Why exactly does rational choice require risk aversion? Debate about this matter is unresolved. So the argument from the veil of ignorance to maximin is highly contentious.

Just as contentious is whether maximin is too strong to be plausible. Suppose you face a choice between directing very small benefits to the worst off (e.g., an additional few hours of healthy life) or directly very large benefits to those a little better off already (e.g., ten additional years of healthy life). Suppose that these are your only options. Here, it seems hard to believe that fairness insists on the comparatively tiny benefits for the worst off over very large benefits for the better off.

In the face of such objections, we might try taking the hard edge out of maximin. A view called *weighted prioritarianism* gives some degree of priority to benefits to the worse off over benefits to the better off, without going so far as to give overriding priority to benefits to the worse off.[[10]](#footnote-8) How much more does a benefit to the worse off count than the same size benefit to the better off? Answers are varied and vague. Nevertheless, the idea does seem deeply attractive.

To see what is attractive about it, suppose we are evaluating two possible distributions. In the first of these distributions, there is great aggregate utility but also a huge gap between the better off and the worst off. In the other distribution, there is a bit less aggregate utility but the worst off is much better off than the worst off are in the first distribution and the gap between the worst off and the best off is much less.

What is the best possible explanation of the fact that the second distribution is morally better? This explanation cannot be that the second distribution has greater utility, because it doesn’t. And the best explanation cannot be merely that equality of welfare is higher in the second, since equality of welfare is not always desirable (e.g., where equality of welfare can be achieved only by levelling down the welfare of the better off to that of the worst off). And the best explanation cannot be supplied by maximin, since that principle is not reliably right in its implications (as we have just seen). The best possible explanation of why a distribution with lower aggregate utility but greater benefits for the worst off might well be that benefits for the worse off simply matter somewhat more than the same size benefits for the better off.

So far, I have outlined five different ideas about what makes rules and social practices substantively fair: (1) the idea that rules and social practices are fair when they maximize expected aggregate welfare calculated in such a way that each counts equally, (2) the idea that rules and social practices are fair when they require reciprocity, (3) the idea that rules and social practices are fair when they ensure, as far as this is possible, that everyone gets what they need, (4) the idea that rules and social practices are fair when they maximin, i.e., leave the worst off as well off as they would be under the best of the alternative rules and practices, (5) the idea that rules and social practices are fair when they maximize expected value where this is calculated in a weighted prioritarian way, i.e., by giving some degree of extra weight to benefits for the worst off. I can now add (6) the idea the social practices of a society are fair if and only if they maximize the extent to which people get what they deserve.

**Section 4: Conflicts between utilitarianism and formal fairness**

Formal fairness and act-utilitarianism can conflict. Suppose a rule is promulgated that medicines will be distributed to those who need them. The first ill person is given medicine. The second ill person is given medicine. The third ill person, however, is not given medicine, because he will in fact waste it, to his own detriment. So we don’t here have formal fairness concerning the rule that medicines will be distributed to those who need them. And this infringement of formal fairness concerning the rule that medicines will be distributed to those who need them would in fact maximize welfare if the medicine that could have been given to the third person were instead given to some other ill person who would benefit from it.

One reaction to this example might be to say that of course the equal and impartial application of any rule other than “choose whatever act maximizes expected utility” can get in the way of maximizing utility. The problem might not really be a conflict between formal fairness and act-utilitarianism but rather a conflict between act-utilitarianism and any other rules.After all, there is no necessity of conflict between formal fairness and the rule “choose whatever act maximizes expected utility.”

That reaction is too glib. Even act-utilitarians admit that constantly making decisions by trying to apply their act-utilitarian principle won’t maximize utility. People often lack information about the probable effects of their choices and, without such information, could not calculate expected value.Furthermore, obtaining such information would often involve greater costs than are at stake in the decision to be made.And, even if people trying to make decisions had the relevant information about possible benefits and harms and their probabilities, calculating expected values is typically unpleasant and time-consuming and thus a cost in itself.And the agent might make mistakes in the calculations. (This is especially likely when the agent’s natural biases intrude, or when the calculations are complex, or when they have to be made in a hurry.)

Finally, there are what we might call expectation effects. Imagine a society in which people know that others are naturally biased towards themselves and towards their loved ones but are trying to make their every moral decision by calculating overall good. In such a society, each person might well fear that others will go around breaking promises, stealing, lying, and even assaulting whenever they convinced themselves that such acts would produce the greatest overall good. In such a society, people would not feel they could trust one another.

For all of the reasons above, even philosophers who espouse the act-utilitarian criterion of moral wrongness reject an act-utilitarian decision procedure. In its place, they typically advocate that, at least normally, agents should decide what to do by applying rules such as “Don’t harm innocent others,” “Don’t steal or vandalize others’ property,” “Don’t break your promises,” “Don’t lie,” “Pay special attention to the needs of your family and friends,” “Do good for others generally.” Hence it is worthwhile to point out that formal fairness in the application of the rules that act-utilitarianism tells us to use as our regular decision procedure can conflict with act-utilitarianism.

Rule-utilitarianism holds that the connection between rules and moral rightness is much tighter than act-utilitarianism holds that it is. Act-utilitarianism accepts that agents should run their daily lives by applying the familiar rules just mentioned, but act-utilitarianism holds that these rules play no part in determining what is morally required, morally permissible, or morally wrong. So act-utilitarianism certainly does hold that which acts are morally required are often not the ones that would be selected by the equal and impartial application of the rules just mentioned. Rule-utilitarianism, in contrast, holds that what is morally required, morally permissible, or morally wrong is determined by whatever rules are the ones whose acceptance would maximize expected utility. Because of the much more central role that rule-utilitarianism accords to rules in determining moral verdicts, rule-utilitarianism will conflict far less than act-utilitarianism with the equal and impartial application of rules such as “Don’t harm innocent others,” “Don’t steal or vandalize others’ property,” “Don’t break your promises,” “Don’t lie,” “Pay special attention to the needs of your family and friends,” “Do good for others generally.”

**Section 4: Conflicts between utilitarianism and substantive fairness**

As I indicated, there are a variety of different ideas about what grounds substantive fairness.

The first one I listed was that the social practices of a society are fair if and only if the social practices have maximum expected utility. The social practices with the greatest expected utility might sometimes require actions with lower expected utility than alternatives. This is why act-utilitarianism and rule-utilitarianism do not always agree about which acts are required, which are permissible, and which are wrong. For, in such cases, act-utilitarianism will of course require the act with the highest expected utility, while rule-utilitarianism will require conformity with the social practice that has the highest expected utility. So, if substantive fairness is determined by the social practices utilitarianism favors, then, since rule-utilitarianism will require conformity with these practices, there will be no conflict between rule-utilitarianism and fairness, even when these social practices require actions with lower expected utility than alternatives. But there will be conflict between act-utilitarianism and fairness in these cases.

The second conception of substantive fairness I listed was the view that social practices of a society are fair if and only if the social practices require each to contribute equally to public goods as long as others are contributing. David Lyons argued that this conception of substantive fairness conflicts with both act-utilitarianism and rule-utilitarianism, though under different conditions.[[11]](#footnote-9)

Suppose enough others are paying their taxes, and I can get away with not paying mine without threat to the public good and without getting punished. My own welfare will be greater if I do not pay taxes, and my paying my taxes would not increase aggregate welfare. In such a case, act-utilitarianism would apparently tell me to keep the money for myself rather than pay my taxes. But fairness definitely requires that, since others are paying their taxes, I should pay mine too. So the objection is that act-utilitarianism cannot explain what is wrong with one person’s free riding on the good behavior of others when this person’s free riding costs no one anything.

There is some controversy about whether this argument is correct about what act-utilitarianism requires. Derek Parfit, for example, argued that, if we avoid various common mistakes in moral mathematics, we will reach the conclusion that the way to maximize expected utility is not to free ride on the contributions of others but instead to do our part.[[12]](#footnote-10) But whether Parfit’s arguments here are sound has been controversial.[[13]](#footnote-11) I think these arguments of his have also received less attention than most of his other work and that the reason for this is that the objection that act-utilitarianism can prescribe not contributing to a public good when there are enough contributions already is not as compelling as many of the other objections to act-utilitarianism (such as objections about harming innocent others for the sake of greater aggregate utility and about the relentlessness of act-utilitarian demands of self-sacrifice for the sake of others).

Unlike act-utilitarianism, rule-utilitarianism can explain why I should comply with optimific rules, even when my compliance will be costly to myself and not beneficial to anyone else. As Richard Brandt wrote,

there could hardly be a public rule permitting people to shirk while a sufficient number of others work. … It would be all too easy for most people to believe that a sufficient number of others were working…. Would it even be a good idea to have a rule to the effect that if one absolutely knows that enough others are working, one may shirk? This seems highly doubtful.[[14]](#footnote-12)

The problem cases for rule-utilitarianism are ones where most others are not following the optimific rules.[[15]](#footnote-13) In these cases too, my complying with the optimific rules would be costly to me but not beneficial to anyone else. So again act-utilitarianism tells me not to follow them. But, while in the earlier cases where others *are* complying with the optimific rules it seems morally required for me to comply with them too, in the present cases where others are *not* complying with the optimific rules it seems morally permissible for me not to comply with them too. So the objection is that rule-utilitarianism tells me to comply with the rules whose widespread acceptance would maximize expected utility, whether or not others are in fact complying with those rules, that is, whether or not others are reciprocating my compliance.

In cases where my complying with a rule would benefit the very people who are not themselves complying with it, then an insistence that I nevertheless comply would threaten to turn me into a “sucker” ripe for exploitation by the “cheats.”[[16]](#footnote-14) Such a requirement on compliers in a world of non-compliance would undeniably be unfair on the compliers and advantageous to the noncompliers.

It would be counterproductive to require people to comply with rules where this compliance is beneficial to those who refuse to follow these rules. One of the best ways of discouraging other people’s noncompliance is to make one’s complying with such rules contingent on their reciprocal compliance. Rule-utilitarianism would thus advocate a rule allowing one to deter the non-compliance of others by refusing to comply unless they do so as well.

But what about cases where noncompliers would not be the beneficiaries of one’s compliance with rules? Cases in which people can make contributions to responsible rescue agencies are good examples. Suppose I believe that if all accepted a rule requiring the comfortably off to give at least 2% of their income to responsible rescue agencies, expected utility would be maximized. But suppose I know that most other comfortably off people won’t give anything like this much. If I nevertheless give at least 2% of my income, the benefit won’t go to the comfortably off people who aren’t doing their share; the benefit will go to those needing rescue. Requiring that I give at least 2% of my income is not unfair on me and does not unfairly benefit those who are in a position to comply with this rule but are not.

I have argued elsewhere that rule-utilitarianism would not tell agents to ignore whether others are doing as they ought. Rule-utilitarianism would instead tell agents to be aware of whether others are doing as they ought—and if they aren’t, then to be willing to make additional sacrifices to protect potential victims from others’ noncompliance.[[17]](#footnote-15) The benefits of having people be willing to do extra to make up to some extent for the noncompliance or underperformance of others are obvious. More people will be rescued than would be if people weren’t willing to make up to some extent for the noncompliance or underperformance of others.

It might seem unfair on those who make their fair share contributions to be required to contribute more than their fair share to make up for others who equally could contribute but aren’t doing so. This is the line of thought developed in Liam Murphy’s *Moral Demands in Nonideal Theory*.[[18]](#footnote-16) Perhaps rule-utilitarians can provide a rationale for a rule that initially require each who is equally able to contribute to contribute an equal amount, which is the person’s fair share contribution. The rationale comes from the diminishing marginal utility of material goods and other resources, as I will now explain.

Normally, the more food or clothes or money or energy people have, the less they benefit from gaining an additional unit. There are many exceptions to this generality. For example, that additional unit of money might put you over some important threshold, such as enabling you to buy something you need. Or that additional bit of energy might enable to surge past the other competitors and win the race. But, as a generality, it is true that material goods and other resources have diminishing marginal utility.

Now if you and I are equally comfortably off, then there is a presumptive case for utility’s being maximized by your and my giving equal shares to responsible rescue agencies. Suppose I give nothing rather than my equal share, and so you give an additional amount to make up for my failure. Other things being equal, the utility loss to you of the additional contribution will be more than the loss to me would have been of my giving my equal share.

However, as I indicated, there are obvious benefits of having people be willing to make additional contributions to rescue agencies to make up to at least some extent for the noncompliance of others. So there are utilitarian reasons to favor a rule requiring people to make up to at least some extent for the noncompliance of others in such cases. There may also be reasons of fairness to require contributions beyond one’s fair share when others aren’t giving their fair share.

Perhaps even more fundamental than the concept of fair share contributions to collective projects (such as funding rescue agencies) is the idea that fairness calls for society to be configured in ways that ensure that, to whatever extent possible, everyone gets what he or she really needs. The idea is that it is unfair for some people to have far more than they need while others have too little to meet essential needs. So if I have more than I need and others have too little to meet their essential needs, then perhaps there are reasons of fairness to require redistributions from me to those without enough to meet their essential needs. These reasons of fairness have to do with my relations to those in need, not with my relations to others who are not in need but in just as good a position as I am to help the needy. So there are reasons of fairness for me to make contributions to help the needy even if others are not giving their fair share.

The idea that fairness requires the satisfaction of needs where this is possible, if this idea is to have even initial plausibility, must rely on a distinction between what is necessary to avoid death, disability, social exclusion, etc., and what is instrumental to the fulfillment of merely optional preferences. You need about half 400,000 calories over the next year to avoid staving to death. You don’t need a straighter nose, whiter teeth, or a more luxurious car, though of course you might prefer all these things. If there are enough resources in society for everyone to have what they need to avoid death, disability, social exclusion, etc., then *perhaps* fairness requires that you get the calories you need to survive. But fairness certainly does not insist that you be given the resources necessary to get your nose straightened, your teeth whitened, and your car upgraded.

Obviously, an account is needed of the distinction between *merely optional* needs and ones that are *not merely optional* *but instead essential or absolute*. A familiar way of drawing this distinction is to say that your merely optional needs are *not* ones that must be fulfilled in order for you to harm but your essential needs *are* ones that must be fulfilled in order for you to avoid harm.[[19]](#footnote-17)

Here “avoid harm” must mean something other than “be worse off than you were before.” The reason is that, because you lack X, you can be worse off than you were before you lacked X and yet you might really need X and are not harmed by lacking X. You might have been better off, for example, before your face started getting wrinkled. But it would be hyperbolic to say that you *needed* a miracle prevention of wrinkles before you started getting them or that you were *harmed* by going from smooth to wrinkled.[[20]](#footnote-18)

Likewise, for the purposes of defining the relevant sense of “need,” “avoid harm” cannot be defined as “be worse off that you would have been otherwise.” The reason is that, because you lacked X, you might be worse off than you would have been otherwise and yet you might not have really needed X and are not harmed by lacking X. For example, you might be worse off now than you would have been had salaries not been frozen, and yet the freezing of salaries was not something that you actually needed not to happen or something that harmed you.

When “need” is defined as that which someone must have in order to avoid harm, I think “harm” needs to be defined as a lack of some essential ingredient of a decent life.[[21]](#footnote-19) For example, you are harmed if your health goes. In trying to give other examples, we quickly get into controversy. We might try contending that you are harmed if you lack access to social engagement, for instance. But what if you don’t want social engagement or even the possibility of social engagement? How can you be harmed by the removal of something you really don’t want anyway?

As popular as the view is that fairness requires that people get what they need insofar as this is possible, the concept of need is too contested and vague, mainly because where the threshold is of a decent life is too contested and vague.[[22]](#footnote-20) But if we are to try to explain fairness *without* any reference to needs, is there a related concept we should employ?

I explained above why I think weighted prioritarianism is a more promising idea than the idea that we should always feel under some moral pressure to equalize welfare or resources, since sometimes the only way to equalize is to level down. I also explained why weighted prioritarianism is a more intuitively appealing than maximin, since maximin would sacrifice huge benefits for the better off for the tiniest gain to those worse off. So we should now ask whether there is indeed a tight connection between weighted prioritarianism and fairness. Does fairness demand that an assessment of benefits and harms accord more importance to benefits and harms to the worse off than it does to the same size (or even larger) benefits and harms to the better off?

The diminishing marginal utility of material goods and other resources (such as energy) militates in favour of redistribution from those who have more material goods and other resources to those who have less. This is a point widely emphasized by utilitarians. But weighted prioritarianism is a view about how gains and losses to utility itself should be assessed. Weighted prioritarianism is not a view merely about how the means to utility (material goods and other resources) get used in utility’s production. So the utilitarians’ point that material goods and other resources typically have diminishing marginal utility must not be conflated with the contention that weighted prioritarians make that gains to the utility of the worse off matter more than smaller, the same size, or even a bit bigger gains to the better off.

There is one very powerful argument in favor of weighted prioritarianism.[[23]](#footnote-21) As I suggested above, that weighted prioritarianism seems to provide the best explanation of why a somewhat smaller but more equal distribution of utility is morally superior to a somewhat larger but less equal distribution. If asked why it is morally superior, we might well say the somewhat smaller but more equal distribution of utility is fairer. And we might then say that what makes the more equal distribution fairer is that it is the one weighted prioritarianism would favor.

Arguments back and forth between utilitarianism and weighted prioritarianism (and between them and maximin) presume that there is no claim on benefits prior to, or weightier than, the size of the benefits, or how many people would get them, or the welfare levels of the parties relative to one anther. But often there is a claim prior to and weightier than such considerations, namely the claim that one person *deserves* some good more than others do. If Jill deserves money or honor or privilege more than Jack does, then she ought to get it, even if he would benefit more from getting it, or even if he is worse off than she is. Recognizing the importance of this sort of consideration, very many people think that the social practices of a society are fair if and only if people get what they deserve.

There are many questions to be raised about desert. How can you deserve what you use your abilities to achieve when you did not deserve the genetic or environmental conditions which fostered the development of those abilities? Is effort the basis of desert or is productivity or is something else? Is one person’s desert always only *comparative* to other people’s desert, or is there *noncomparative* desert?[[24]](#footnote-22)

The most important question about desert seems to me to be whether all principles of desert are “institutional” or whether some are “pre-institutional.” John Rawls argued that the institutional account of desert understands all principles of desert as entirely derivative from of *just* institutions.[[25]](#footnote-23) On this view, just institutions must be determined by grounds other than desert. The pre-institutional account of desert, in contrast, holds that, while many principles of desert make no sense without the embedding of an institutional context, some principles of desert do *not* depend the selection of institutions and social practices on grounds other than that they give people what they deserve.

On the debate between institutional and pre-institutional accounts of desert, utilitarians agree with Rawls. Utilitarians evaluate institutions and social practices purely by how much utility would result from their establishment, not by whether they would give people what they deserve. And those who advocate a maximin criterion evaluate institutions and social practices by the results for the worst off. Weighted prioritarians assess institutions and social practices in terms of the aggregate welfare where this is calculated with extra weight given to benefits for the worse off. All these views are versions of institutionalism about desert.

The question of whether all principles of desert are determined by institutions and social practices or instead some principles of desert are pre-institutional seems to me more important than other questions about desert because institutionalism gives us ways of dealing with those other questions. For example, institutionalism enables us to decide on whether effort should be the basis of desert, or productivity, or some combination of them, or some alternative to them. Utilitarian institutionalists, for instance, would decide by asking which reward structure would maximize aggregate welfare calculated by giving the same weight to a benefit to any one individual as to the same size benefit to any other individual. Weighted prioritarians would decide by asking which reward structure would maximize aggregate welfare calculated by giving the somewhat more weight to benefit to the worse off.

Even if utilitarian institutionalists are correct about which institutions and social practices determine the principles of desert, there is room for divergence in particular cases between what rule-utilitarianism requires overall and what people deserve. For rules about desert constitute not the whole but only a subset of the rules that rule-utilitarianism endorses, and in particular cases conflicts between rules about desert and other kinds of rules can arise. Perhaps rules about desert are particularly weighty and so normally trump whatever other rules come into conflict with them. Still, it is hard to believe that rule-utilitarians will think that rules about desert always trump other kinds.

The scope for conflict between desert and utility is far greater with act-utilitarianism. An act-utilitarian has to believe that, even if certain institutions, social practices, and their constituent rules are perfectly well justified on utilitarian grounds, cases regularly arise in which complying with the perfectly well justified rule about desert will predictably not maximize utility. This is true, for example, when unpopular people deserve rewards, or when the people who deserve the rewards will not actually benefit much from them.

**Section 5: Conclusion**

This chapter has distinguished between the main kinds of utilitarianism—act-utilitarianism and rule-utilitarianism. It has also distinguished between different kinds of fairness.

Formal fairness consists in the equal and impartial application of rules. With respect to the question of which acts are morally required, act-utilitarianism regularly offers a different verdict from the one implied by application of rules other than the rule “always choose the act that will maximize utility.” In this way, the equal and impartial application of rules other than “always choose the act that will maximize utility” will conflict with act-utilitarianism.

Substantive fairness is a matter of the content of the rules, whether the rules make all and only the distinctions that fairness requires. Fairness requires that a distinction be made between those who are reciprocating or are willing to reciprocate kindness towards them and those who are not willing to do this. Is the application of this distinction one that rule-utilitarianism can endorse? I argued that it is. But the application of this distinction conflicts with act-utilitarianism in many cases.

Fairness seems to consist of more than merely formal fairness plus reciprocity. One popular idea of fairness is that it is unfair for society to be arranged in such a way that some people end up with more than they need while others have less than they need. Where is the threshold between what is needed and what is good but not needed? Despairing of a persuasive answer to that question, I offered weighted prioritarianism as alternative to thinking in terms of needs. Weighted prioritarianism is a rival to utiltiarianism, and some objections to utilitarianism on the grounds of fairness are not also objections to weighted prioritarianism.

The concept of desert is perhaps even more central than the concept of need to popular thinking about fairness. There are many questions about the concept of desert. I suggested that the most important of these is whether all the principles of desert are derivative from institutions, social practices, and their constituent rules that are justified on grounds that do not include principles of desert, or instead whether there are some principles of desert that are “pre-institutional.” If all the principles of desert are derivative from institutions, social practices, and their constituent rules that are justified on grounds that do not include principles of desert, then thinking about which possible institutions and rules would be best on those grounds might help us with those other questions about desert. However, even utilitarian institutionalists about desert will admit that there is scope for conflict between desert and utilitarianism. Again, possible conflict between rule-utilitarianism and desert will be less than between act-utilitarianism and desert.

1. Satisficing act-utilitarianism is drawn from Michael Slote. See his “Satisficing Consequentialism,” *Proceedings of the Aristotelian Society* supplementary vol. 58: 139–63; *Common-Sense Morality and Consequentialism* (Oxford: Oxford University Press, 1985); *Beyond Optimizing* (Cambridge, MA: Harvard University Press, 1989); and *From Morality to Virtue* (New York: Oxford University Press, 1992). [↑](#footnote-ref--1)
2. Peter Vallentyne, “Against Maximizing Act Consequentialism,” in James Dreier (ed.), *Contemporary Debates in Moral Theory*, Boston, MA: Blackwell Publishers, pp. 21–37, esp. pp. 23–8. [↑](#footnote-ref-0)
3. Tim Mulgan, *The Demands of Consequentialism*, Oxford: Oxford University Press, pp. 129–42. [↑](#footnote-ref-1)
4. See Slote, *Commonsense Morality and Consequentialism*, ch. 5; Alasdair Norcross, “Scalar Act-utilitarianism,” in Henry West (ed.), *Blackwell Guide to Mill’s Utilitarianism*, Boston: Blackwell Publishers, pp. 217–32. [↑](#footnote-ref-2)
5. Philip Pettit and Michael Smith, ‘Global Consequentialism’, in Brad Hooker, Elinor Mason, and Dale E. Miller (eds.) *Morality, Rules and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press, 2000), pp. 121–33; and Shelly Kagan, ‘Evaluative Focal Points’, in Hooker, Mason, and Miller (eds.), pp. 134–55. [↑](#footnote-ref-3)
6. Henry Sidgwick, *Methods of Ethics*, 7th edition, London: Macmillan, 1907; J. J. C. Smart, “Outline of a System of Utilitarian Ethics,” in Smart and Bernard Williams, *Utilitarianism: For & Against*, Cambridge, Cambridge University Press, 1973, pp. 3–74. [↑](#footnote-ref-4)
7. For fairly recent discussions, see my *Ideal Code, Real World*: *A Rule-consequentialist Theory of Morality*, Oxford: Oxford University Press, 2000; and Tim Mulgan, *Future People*, Oxford: Oxford University Press, 2004. [↑](#footnote-ref-5)
8. See John Rawls, “Justice as Fairness,” *Philosophical Review* 67 (1958), pp. 164–9, p. 164, n. 2. [↑](#footnote-ref-6)
9. Most influential was his *A Theory of Justice*, Cambridge, MA: 1971. [↑](#footnote-ref-7)
10. See Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986), p. 227; Thomas Nagel, *Equality and Partiality* (New York: Oxford University Press, 1991), ch. 7; Derek Parfit, “Equality and Priority,” *Ratio* 10 (1997), pp. 202–21; David Miller, *Principles of Social Justice* (Cambridge, MA: Harvard University Press, 1999), pp. 223–5; Marc Fleurbaey, Bertil Tungodden, and Peter Vallentyne, “On the Possibility of Nonaggregative Priority for the Worst Off,” *Social Philosophy and Policy* 26 (2009), pp. 258–85. [↑](#footnote-ref-8)
11. See David Lyons’s *Forms and Limits of Utilitarianism*, Oxford: Oxford University Press, 1965, pp. 139–41. [↑](#footnote-ref-9)
12. Derek Parfit, *Reasons and Persons*, Oxford: Oxford University Press, 1984, ch. 3. [↑](#footnote-ref-10)
13. See James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance*, Oxford: Oxford University Press, 1986, pp. 215–219; Kristin Schrader-Frachette, “Parfit and Mistakes in Moral Mathematics,” *Ethics* 98 (1987), pp. 50–60; Bjon Petersson, “The Second Mistake in Moral Mathematics Is Not About the Worth of Mere Participation,” *Utilitas* 16, pp 288–315; Ben Eggleston, “Does Participation Matter? An Inconsistency in Parfit’s Moral Mathematics,” *Utilitas* 15 (2003), pp. 92–105. [↑](#footnote-ref-11)
14. Richard Brandt, “Some Merits of One Form of Rule Utilitarianism,” *University of Colorado Studies in Philosophy*, 1967, pp. 39–65, fn. 15. [↑](#footnote-ref-12)
15. Lyons, pp. 128–32, 137–42. [↑](#footnote-ref-13)
16. See J. L. Mackie, “The Law of the Jungle: Moral Alternatives and the Principles of Evolution,” *Philosophy* 53 (1978), pp. 455–64; “Co-operation, Competition, and Moral Philosophy,” in A. M. Colman (ed.), *Cooperation and Competition in Humans and Animals* (Wokingham: van Nostrand Reinhold). [↑](#footnote-ref-14)
17. See my *Ideal Code, Real World*, pp. 164–9. [↑](#footnote-ref-15)
18. Liam Murphy, *Moral Demands in Nonideal Theory*, New York: Oxford University Press, 2000. [↑](#footnote-ref-16)
19. Joel Feinberg, *Social Philosophy*, Englewood Cliffs, NJ: Prentice-Hall, 1973, p. 111; David Wiggins, “The Concept of Need,” in Wiggins, *Needs, Value, Truth*, 2nd edition, Oxford: Blackwell Publishers, 1991, p. 10. [↑](#footnote-ref-17)
20. If getting wrinkled caused you to lose your spouse or your job, things would be different. I am assuming that you don’t have the sort of spouse or job that you would lose because of wrinkles. [↑](#footnote-ref-18)
21. An influential advocate of the view that needs are what one must have satisfied if one is to have a decent life has been David Miller. See his *Principles of Social Justice*, Cambridge, MA: Harvard University Press, 1999, especially pp. 212, 319 n. 23, n. 25. [↑](#footnote-ref-19)
22. See my “Fairness, Needs, Desert,” in Matthew Kramer, Claire Grant, Ben Colburn, and Antony Hatzistavrou (eds.) *The Legacy of H. L. A. Hart*, Oxford University Press, 2008, pp. 181–99. [↑](#footnote-ref-20)
23. I discuss powerful arguments *against* weighted prioritarianism in *Ideal Code, Real World*, pp. 60–65 and in ‘When Is Impartiality Morally Appropriate?’ in Brian Feltham and John Cottingham (eds.), *Partiality and Impartiality: Morality, Special Relationships and the Wider World* (Oxford: Oxford University Press, 2010), pp. 26–41, at p. 39. [↑](#footnote-ref-21)
24. For relevant discussions, see the essays collected in Louis Pojman and Owen McLeod (eds) *What Do We Deserve?* New York: Oxford University Press, 1999. See also Shelly Kagan, *Desert*, Oxford: Oxford University Press, 2012. [↑](#footnote-ref-22)
25. John Rawls, *A Theory of Justice*, §§ 17 and 48. For a helpful discussion, see Serena Olsaretti, “Distributive Justice and Compensatory Desert,” in Olsaretti (ed.) *Desert and Justice*, Oxford: Oxford University Press, pp. 187–204, at pp. 196–7. [↑](#footnote-ref-23)