CHAPTER SEVEN

IS UTILITARIANISM TOO DEMANDING?

It is clear that since for the utilitarian only the best is good enough, there is no room for doing less good than you can. If you fail to do what is best in terms of total well-being, you are simply doing wrong. But maximizing total well-being can be very demanding, as the following illustrations show.

HEROISM

After you have saved a child from a burning building you are told that there is another child still left in the building. You could go back into the building and save the second child, but you know that this will cause you third-degree burns. According to utilitarianism, this would not just be a heroic thing to do; you *ought* to do it, and it is wrong not to do it.

YOUR MONEY AND CHARITY

You are wondering whether to spend a pound on chocolate for yourself or to give it to a certain charity. You know that this charity is unusually effective and that even a small contribution can help them save a child from some crippling and painful illness. Since you obviously do more good by saving a child from illness than by eating a piece of chocolate, you ought to give the pound to charity. However, if you repeat this utilitarian reasoning every time you have a pound to spare, you will end up very poor indeed.

EVERYDAY LIFE

When you, after a long and tiring day, put on your slippers and watch Celebrity Big Brother on TV, you are probably acting wrongly.

There will almost always be opportunities to produce greater good: meet a lonely relative, talk to your depressed neighbour, do some charity work etc. (I ignore the possibility that the TV show is so bad that it is bad for you to watch it.) Is it sensible to say that you violate your moral duty whenever you spend a quiet evening at home in front of the TV? This would be a heavy duty indeed. We are all wrongdoers when we are relaxing.

In these cases, we seem to think that to maximize goodness is beyond the call of duty. It is a morally desirable thing to do but it is not something you ought to do. Another way to sum up the problem is to say that utilitarianism does not give people *options*; it does not allow the agent to pursue his interests at the expense of the overall good.

In this chapter I will first go through some of the standard utilitarian responses to these cases. I will then consider how demanding utilitarianism is in comparison to other non-utilitarian moral theories. In the final section, I will discuss the objection that utilitarianism is so demanding that it cannot be reconciled with true friendship.

UTILITARIAN RESPONSES

(a) Ought entails can

The utilitarian could try to lessen the blow by pointing out that only heroes can perform heroic acts. Normal people do not have the guts and strength to perform heroic actions. For instance, paralyzed by fear, you might not be able to go back into the burning building to save the second child. But if this is so, then you are not obligated to act heroically, for ought entails can.

One obvious reply is to say that even if you are unable, in your present state, to perform heroic actions, you could still change yourself gradually over time and acquire the ability to perform heroic actions. However, this assumes a very optimistic picture of humans. It is doubtful that we can all become moral saints given the right training. Many of us would just make things worse if we tried. The pessimistic utilitarian would therefore say that we ought not even try to become saints.

No matter whether this pessimism is warranted, the 'ought entails can' reply can in any case only provide a partial answer, since there are many cases, the charity case above being one example, where it is

clearly possible for us to sacrifice a lot in our lives for the sake of the overall good.

(b) Distinguish wrongness from blameworthiness

A more general response that seems to take most of the sting out of the objection is to deny that wrongness entails blameworthiness. The utilitarian claim that we are almost always doing wrong is especially hard to accept if one assumes that all wrongdoers are blameworthy. Now, it is easy to deny that wrongness entails blameworthiness if blameworthiness is understood as 'ought to be blamed' and blaming is, in turn, understood as an act of 'telling someone off'. It is clear that utilitarianism does not say that we should always tell wrongdoers off, for telling a wrongdoer off is an action that will often only make things worse.

But this understanding of blameworthiness is questionable. To judge that someone is blameworthy seems more like an assessment of *him* than an assessment of an *act* directed towards him. Indeed, it seems possible for someone to be blameworthy even though no one can 'tell him off'. Hitler is still blameworthy, but since he is dead we can no longer tell him off. Similarly, an evil person who has fallen into a coma can still be blameworthy even if no one can tell him off (and get the message across).

This complication need not worry the utilitarian, however, because he simply can take it on board: To say that an act is wrong is to assess the act, not the agent. So, when the utilitarian is saying that you do wrong when you relax in front of the TV, he is not saying that you are blameworthy, for to say that an agent is blameworthy is to assess the agent, not the act.

But this reaction is in fact too hasty. Even if wrongness and blameworthiness are different concepts that apply to different things, there can still be necessary connections between these concepts. In particular, we have not yet ruled out that if I do something wrong, I must be blameworthy. We need a positive argument against such a connection.

One such argument has to do with moral excuses. Sometimes we have a good excuse for why we acted wrongly. But if we have a good excuse, we are not blameworthy. Suppose that, walking down the street one day, you see a person apparently stabbing another with a knife, and you try to stop this by hitting the attacker. Now, the apparent

assault was in fact only play-acting in a street theatre performance, but you could not have known this at the time. Your injuring the actor seems wrong, but since you had no clue that it was only a staged attack, you are not blameworthy.[1]

Of course, there are still cases where we do seem to know that we are not maximizing the good and thus know that we are doing wrong according to utilitarianism. For example, when you are relaxing in front of the TV you may in fact know that you could do more good by calling a friend in distress or help your neighbour with babysitting. So, the utilitarian is pressed to offer a more general response.

One such response would be to say that even if we concede that knowingly doing wrong merits blame, how much blame it merits depends in part on how costly the action would be to the agent. If helping your friend would only require a quick phone call, then not calling your friend merits more blame than it would in a situation where helping your friend would require spending a whole day and night with him. Similarly, if you let your friends die when you could have saved them by sacrificing your life, you have done wrong, but since the costs to you of doing right would have been enormous, you do not merit much blame. Indeed, one might even think that you do not merit any blame for your wrongdoing, since you have a very good excuse.

(c) Morality is demanding but not overriding

The utilitarian could ask us whether we really know the reasons for our worry that utilitarianism is too demanding. Are we sure it is a *moral* worry? Do we question utilitarianism because of its demanding moral requirements, or because these moral requirements are supposed to override all other non-moral requirements, including requirements of prudence, friendship, and parenthood? Perhaps the crux of the matter is that we implicitly assume that utilitarianism tells us that *all things considered* we ought to sacrifice a lot in our lives. But this follows only if we accept:

> *Overridingness*
> In deciding what to do *all things considered,* moral reasons overrides any other kind of reason.

If this thesis is denied, the utilitarians could say that it is true that morality is very demanding but there are non-moral reasons that sometimes override moral reasons. For instance, even if morality requires you to sacrifice your life in order to save your friends, prudence might override morality in this case so that you are not required, all things considered, to sacrifice your life. A utilitarian who denies the overridingness thesis can therefore claim that your all things considered requirement will not be too demanding.

However, the overridingness thesis seems compelling, for what is the all-things-considered ought if not the moral ought?

(d) Reject maximizing, accept satisficing

One reason why utilitarianism is so demanding is that it tells us to maximize value, to do the best you can. The right remedy might thus be to lower the standards and say that you are only obligated to do what is *sufficiently* good, not what is best:

> *Satisficing utilitarianism*
> An action ought to be done if and only if it would bring about a sufficient level of total well-being.[2]

To make this theory more precise, we need to decide on a non-maximal level of total well-being, *w*, that counts as sufficient in the circumstances. This theory might look promising, since it allows you to choose between the actions that will produce a sufficient amount of good (at least as much as *w*), and thus permits you to pursue your interests at the expense of some overall good.

But satisficing utilitarianism will in fact not solve the problem with heavy duty, for the agent is still required to produce an outcome with the overall well-being of at least *w*, no matter how much this will cost the agent. So, for instance, if the agent can produce an outcome with *w* amount of total well-being only if he makes a big sacrifice, he is required to do so. The root of the problem is that *w* is still defined as a sum of total well-being. So, in order to avoid this problem, we have to somehow single out the *agent's* sacrifice and give it extra weight when we define the sufficient level of total well-being. But this runs counter to impartiality – the agent's well-being should be given the same weight as the well-being of any other person – and so is not an option for utilitarians.

(e) Utilitarianism is not as demanding as we might think

Contrary to appearance, utilitarianism does not imply that you should walk through life like a pained do-gooder constantly trying to save people from illnesses and death. There are three reasons for this.

First, you will be burnt out if you are constantly trying to promote well-being. Charity workers often complain about how hard it is to be surrounded by ill and dying people. Even Mother Theresa, who is often seen as a moral saint, admitted in her diaries that she often suffered from depression. Now, giving yourself some periods of relaxation will not just do you some good but will also do other people good, since it will make you a more efficient promoter of total well-being. We tend to miss this obvious fact because we think about our options as immediate one-shot actions: Should I *now* save this child from illness? Should I *now* donate one pound to charity? But our options often include plans of actions that stretch into the future. So, the question is not whether I should now donate this pound or save this child; the question is whether I should include in my plan for the future a certain amount of charity work. If I do not leave any time for relaxation in my plan, I will be exhausted and do less good overall. The best plan available to me will contain an optimal balance of periods of relaxation – 'me' time – and periods of charity work. How this balance will look may differ from one person to another depending on skills, motivation, and knowledge.

Second, freelance do-gooders need not be the best promoters of well-being. It is often more efficient to unite and act together to change crucial political and economic structures that prevent poor and ill people from flourishing. For example, instead of sending almost all of your salary to Oxfam every month you should get together with others and put pressure on your government to write off poor countries' international debts.

Third, you may be able to change your character and values so that helping others in need becomes one of your deepest projects. Of course, this is not a change that will happen over night. You can only do it indirectly by implementing a long-term plan, which may involve finding out more about poor people's needs and joining a local charity organization. Now, the more you desire to do good, the less burdensome it will be. You will of course be forced to sacrifice a lot of time, energy, and comfort, but you do not have to sacrifice all of your deepest projects since doing good will now be one of them.

Also, the more you desire to do good things, the more likely it is that you will succeed in doing good, since success depends partly on motivation. Utilitarianism will therefore tell you to work against world poverty *and form an interest in doing it.*

NON-UTILITARIAN DEMANDS

You may not be convinced by these responses. But it is important to keep in mind that non-utilitarian moralities are quite demanding too. For instance, it is not clear that they will avoid the 'chocolate versus charity' problem. It is not just utilitarians who think we should rather save a child from illness than buy a piece of chocolate. Surely, any plausible non-utilitarian theory must accept that it is, other things being equal, better to save a life than to enjoy a piece of chocolate. But then, if this reasoning is repeated, we get the result that the agent becomes poor, since small sacrifices will eventually add up to a big sacrifice.

Some non-utilitarian moralities threaten to be as demanding as utilitarianism. For example, a virtue ethics that tells you to do what the fully virtuous person would do is potentially very demanding. It all depends on how the fully virtuous person is defined. If he is in the league of Gandhi, Jesus, and Mother Theresa, we will have a very demanding theory that tells you to endure great sacrifices for the sake of others.

Of course, the virtue ethicist could respond by saying that the fully virtuous person should be seen as an *ideal* for us to emulate in our actions as much as possible. The closer our actions resemble this ideal, the better they are. But the utilitarian can say something similar. The fully impartial and benevolent agent can also be seen as an ideal. The closer your actions resemble this ideal, that is, the more total well-being you produce, the better your actions are. Now, there is a general worry here that the notions 'right' and 'wrong' would no longer have any role to play. However, if this is a real worry, then it is as much a worry for virtue ethics as it is for utilitarianism. Notice, for instance, that if the virtue ethicist responds to this worry by equating the right action with the ideal, then all actions falling short of the ideal will be wrong, and he would have to agree with the utilitarians that most of us are wrongdoers most of the time. Of course, he could still say that there are degrees of wrongdoing – ethnic cleansing is a more serious wrong than a white lie. It is thus a mistake to think that

nothing matters just because we are all wrongdoers most of the time. There are better and worse sinners. But, again, this option is open to the utilitarian as well.

Kantianism cannot avoid the demandingness problem either, since it imposes a strong prohibition against deception; you are not allowed to lie even if this is the only way to protect your vital interests. Indeed, according to Kant, you are not allowed to lie even if this is the only way to save your friends or loved ones from being killed by the enemy soldier who is asking you where they are hiding.

BEYOND THE CALL OF DUTY

Still, there seems to be one important difference between utilitarian demands and non-utilitarian demands. Utilitarians cannot accept that there are *supererogatory* actions, actions that are morally desirable, but go beyond the call of duty. Non-utilitarians, in contrast, can accept such actions, since they can reject the consequentialist idea that you ought to do the best you can.

It is true that we do often describe actions as going beyond the call of duty, but it is in fact not clear how to make sense of this talk. The supererogatory action is supposed to be better in some sense than the action you ought to perform. But if 'better' means 'more moral reason to do', we have a problem. We would have to say that you are morally permitted to do something even though you have more moral reason to do something else. This sounds paradoxical.

One way to avoid this paradoxical situation is to make use of the distinction between wrongness and blameworthiness we discussed earlier. You *should* or *ought* do the supererogatory action, but you do not *have* to. You ought to do it, because this is what you have most moral reason to do. But you do not have to do it in order to avoid blame. So, although failing to save the second child in the burning building is wrong, it does not make you blameworthy, for you have a very good excuse (third-degree burns). You ought to save her but you do not have to.

This way of avoiding the paradox would of course be welcomed by utilitarians and consequentialists, since they are keen to distinguish wrongness from blameworthiness. If it can be made to work, utilitarians will be able to accommodate supererogatory actions after all.

CAN UTILITARIANS BE GOOD FRIENDS?

Some people think that the answer is clearly no. It is not just that utilitarians do not make good friends; they do not make friends at all. This would be a serious objection to utilitarianism, since loving relationships and friendships are among the most important factors in a good life, a life worth living. So, if utilitarianism requires agents to lead lives that are not worth living, it is surely an all too demanding theory. But is it really true that utilitarians cannot be friends? To answer this question, let us set up a dialogue between a utilitarian and a critic, who used to think they were friends.

The critic: I hope you agree that if you are my genuine friend, you value and love me as an *end*. You value me for what I am over and above the usefulness of being your friend. And you do things for my sake, not just for the sake of some other value. If you value me merely as a useful acquaintance, who can help you out in different ways, then you are not a genuine friend. How can you then be my friend? As a genuine utilitarian, you have only one fundamental aim in your life: to make the world a better place. So there cannot be any room for other commitments in your life. In particular, there can be no room for friendship.

The utilitarian: I agree with your analysis of friendship. But it is a misunderstanding to think that my utilitarian theory prevents me from pursuing other ends and interests. My utilitarian theory does allow me to be your true friend. You fail to distinguish between utilitarianism seen as a *decision method* and utilitarianism seen as a *criterion of rightness*. Utilitarianism is first of all a criterion of rightness. It tells you what makes an action right or wrong. This is not the same thing as a method of deliberation which tells you what you should aim at and how you should deliberate when you decide what to do. Since I believe in utilitarianism, I believe that what makes an action right is that it has optimal consequences in terms of overall well-being. But I do not believe that I should constantly be preoccupied with utilitarian calculations. For this will itself have consequences, and sometimes bad consequences. This applies to friendship as well. I think that the world is a better world when people have relationships like ours and, as you have pointed out, I could not have this relationship if I constantly made utilitarian calculations. I would make things worse if I took an instrumental attitude towards my

friendships and always applied the utilitarian test to them. Of course, this is not to give up utilitarianism. For I still have a standing commitment to lead the life that will have the best consequences overall. I would seek to lead a different sort of life if I did not think that this life promoted overall well-being.

The critic: But how can you value me as an end when you would happily end our friendship if you thought it would not be conducive to overall well-being. Your commitment to utilitarianism will always *override* your commitment to me. This shows that you are not my friend.

The utilitarian: No, you confuse the notion of commitment to an end *for its own sake* with that of an *overriding* commitment. I do value you as an end even though I do not value you as an overriding end. Not all ends are overriding. For instance, I am sure that in a choice between ending our relationship and causing horrible suffering to your kids, you would choose to end our relationship. But this does not show that you are not my friend. Does it?

The critic: OK, you may be right about this. But there is still a contingency involved in your attitude towards me that prevents you from valuing our friendship as an end. Look, you are telling me that

(a) you value me as an end.

But since you are a utilitarian it is also true that

(b) you value me only so long as valuing me promotes overall well-being.

But to value me for my own sake is to value me for what I am in myself, in virtue of who I am. So, (a) is incompatible with (b).

The utilitarian: Your analysis of what it is to value someone for their own sake is correct, but you apply it incorrectly. (b) does not say that I take an instrumental attitude towards you. It says that my attitude is *contingent* on its promoting overall good. It says that my intrinsic attitude towards you meets a certain *counterfactual* condition: I would not hold this attitude if it did not promote the overall good. And there is nothing strange in saying that an intrinsic attitude is contingent in this way. Consider a tennis player who desperately wants to win. It is self-defeating for him to think, 'No matter how

I play, the only thing I care about is whether I win.' He should instead devote himself more to the game and the play as such. Assume that he spends a good deal of time developing a devotion to the play as such so the following is now true of him.

(c) He values playing the game for its own sake.

Then it is not true that he values the game instrumentally. But this does not mean that his initial goal to win is inconsistent with his devotion to the play. For his desire to win can still be *an ultimate organizing desire* for his other desires. If this desire is effective, the following will be true about him.

(d) He values playing the game only so long as valuing the game promotes his winning.

His intrinsic devotion to the play is *contingent* on his winning the games. If this motivational structure is possible in the case of tennis playing, it must be possible in the case of friendship.[3]

The critic: Well, I am not a tennis game, am I? I am a person. I am not fully convinced. Perhaps you can value me as an end. So perhaps you can even be my friend after all. But you are not an especially good friend I must say. You say that you value me for my own sake. But you do not value me *wholeheartedly*. To value someone wholeheartedly is not just to value that person for her own sake but also to intrinsically endorse that very attitude. It is to have an intrinsic positive attitude towards your first attitude. Your first-order attitude requires a second-order intrinsic endorsement. In other words,

(e) If you *wholeheartedly* value me as an end, then you intrinsically endorse the fact that you value me as an end.

Compare with a drug addict who loves the high for its own sake but intrinsically hates the fact that he loves the high. He does not want to be a person who loves the high. This person does not wholeheartedly love his high, for he does not intrinsically endorse his love for the drug. Or take the case with the kleptomaniac who loves the rush he experiences when he shoplifts, but hates the fact that he loves this rush. He does not wholeheartedly value this rush because he does not intrinsically endorse the fact that he likes the rush. It is the same with

you and your attitudes towards me. For you only value your first-order attitude towards me *instrumentally*, only so long as it promotes goodness. So, you cannot value me wholeheartedly. You do not endorse the fact that you value me as an end.

The utilitarian: No this is not right. First of all, your definition of what it is to wholeheartedly value something is too demanding. Look at the case with the drug addict. He hates the fact that he loves the high. But my attitudes towards you are different. I do not hate the fact that I value you as an end. So, the right characterization of a wholehearted engagement is

(f) If I *wholeheartedly* value you as an end, then I do not intrinsically hate the fact that I value you as an end.

Furthermore, my consequentialist commitment does not prevent me from intrinsically endorsing the fact that I value you as an end. For this second-order attitude might be the attitude that promotes the overall good in the long run. I am inclined to think so. The world would be a worse place if I did not intrinsically endorse the fact that I love you as an end. My commitment to you would not be strong enough to promote the good things that come out of an intimate friendship.

You have been questioning me. Let me now ask you something. If you are right in your accusations then all moral agents seem to have a problem no matter which moral theory they believe in. For your criticism can of course be generalized. You seem to think that it is impossible to combine the following two claims:

I value you for your own sake
I value you only so long as this promotes the overall goodness.

But your argument works even if we replace the utilitarian moral goal with a non-utilitarian one:

I value you for your own sake.
I value you only so long as this satisfies the non-utilitarian moral principle P.

For P we can put any non-utilitarian moral principle. For instance, P can be Kant's categorical imperative, or some set of deontological

principles, the Ten Commandments, a virtue theory, or what have you. So if you are right, then moral theory in general is in trouble. You cannot be deeply committed to a moral theory if you want to be a good friend. But that seems to be a reductio of your argument.

The critic: No, this is not true. Now *you* are conflating things. It is true that some moral theories will give moral agents a hard time being friends. But not all. It depends on the contents of these theories. I, for one, think that an acceptable moral principle must explicitly and directly allow me to value people as ends and act on these values. For instance, I should be permitted to spend time with my friends instead of making sure that a bunch of strangers spend time with their friends. Also, I should be permitted to visit my friend at the hospital even though I know that the lonely guy next door would benefit more from a visit. What is wrong with utilitarianism is that it cannot allow these actions. Your theory is too impartial.

The utilitarian: But a moral principle should be impartial. We should give equal considerations to all people. This is what it means to be moral. So I am not convinced. But I think you have at least convinced me that one should not discuss the value of friendship if one wants to keep one's friends. It may be impossible to be friends with a utilitarian who tries to convince you that she can be a good friend.

CONCLUDING REMARKS

There is no way around it, utilitarianism is a demanding moral theory. Utilitarianism implies that we are all moral underachievers, since we often fail to maximize total well-being. However, even if it is true that we often act wrongly, it is not clear that this also means that we are blameworthy, for in many cases, we do have a good excuse for doing wrong (we did not know it was wrong, or the right action would have been very costly to us). But it is doubtful that we have a good excuse for not giving more to charity and for not spending more time and energy on helping the needy. So, utilitarianism will definitely blame us for not being more other-regarding in our actions. Of course, no one doubts that being more other-regarding is praiseworthy and morally desirable, but the question is whether this is what we ought to do.

However, utilitarianism is not alone in being a demanding moral theory. Both virtue ethics and Kantianism can make tough demands

on agents. Even common-sense morality, which is usually supposed to be pretty easy on moral agents, does ask for great sacrifices if reasonable demands are repeated over time. So, it seems that any plausible moral theory will be demanding. I leave it to you to decide whether utilitarianism is too demanding.

SUGGESTED READING

On the demandingness objection:

Cullity, G. (2004), *The Moral Demands of Affluence*, Oxford: Clarendon Press.
Mulgan, T. (2001), *The Demands of Consequentialism*, Oxford: Oxford University Press.
Singer, P. (1972), 'Famine, affluence, and morality', *Philosophy and Public Affairs*, Vol. 1, No. 3, 229–243.

On blameless wrongdoing:

Parfit, D. (1993), *Reasons and Persons*, Oxford: Oxford University Press, Part 1, Chapter 1, Section 14.
Tännsjö, T. (1998), *Hedonistic Utilitarianism*, Edinburgh: Edinburgh University Press, Chapter 9.

On the notion of going beyond the call of duty (supererogation):

Heyd, D. (1982), *Supererogation*: *Its Status in Ethical Theory*, Cambridge: Cambridge University Press.
Zimmerman, M. (1994), *The Concept of Moral Obligation*, New York: Cambridge University Press, Chapter 8.

On friendship and utilitarianism:

Railton, P. (1984), 'Alienation, consequentialism, and the demands of morality', *Philosophy and Public Affairs*, Vol. 13, No. 2, 134–171.
Scanlon, T. M. (1998), *What We Owe to Each Other*, Cambridge, MA: Harvard University Press, 88–89.

CHAPTER EIGHT

IS UTILITARIANISM TOO PERMISSIVE?

We have seen that utilitarianism can be accused of being too demanding since it does not give us any *options*; it does not allow us to pursue our innocent projects at the expense of overall well-being. But utilitarianism can also be accused of being too permissive, because it rejects both *constraints* on actions and *special duties* to our nearest and dearest. It rejects constraints on actions, since any action, no matter how morally repugnant, can be obligatory. We just need to imagine a case in which a repugnant action happens to maximize total well-being. For instance, I am allowed to torture an innocent person if this is the only way for me to promote overall well-being. Utilitarianism rejects special duties as well, since impartial benevolence cannot admit of any exceptions in favour of some people over others. So, for example, I am permitted to save the stranger's child rather than my own, if the stranger's child would benefit more.

In this chapter I shall consider the utilitarian responses to these objections. In particular, I shall give an explanation of why classical utilitarianism is unable to accommodate constraints and special duties. This will enable us to see whether other forms of utilitarianism can do a better job at accounting for constraints and special duties.

CONSTRAINTS

Examples of constraints embraced by common-sense morality as well as many deontological theories are,

Do not lie!
Keep your promises!

Do not kill innocent persons!
Do not torture!
Do not punish the innocent!

All deontologists agree that constraints cannot easily be overridden by considerations about well-being. You are forbidden to violate them even if this is the only way to promote total well-being. Are you *never* allowed to violate them? Here deontologists differ in their views. Some think these constraints are absolute so that you are never, under any possible circumstances, allowed to violate a constraint. Others think a constraint can be violated when it comes into conflict with a more important constraint. Perhaps you are allowed to lie in order to avoid breaking an important promise. Still others go even further and think that a constraint can be violated when enough total well-being is at stake. Perhaps you are allowed to kill one innocent person in order to prevent a nuclear holocaust.

Since utilitarianism does not accept any constraints on actions (except, trivially, the constraint to maximize total well-being), it can be accused of being too permissive. Here are some vivid examples, listed in order of increasing repugnancy.

The promise

You have promised to return a book to your friend. But you realize that your lonely neighbour would benefit a lot from reading it. So, you decide to give the book to your neighbour instead of returning it to your friend. Utilitarianism would approve of your action if you benefited your neighbour more than you harmed your friend.[1]

The car accident

One winter, you have had a car accident on a lonely road. The other passenger is badly injured. You find an isolated house occupied by an old woman and her grandchild. There is no phone, but a car in the garage. You ask to borrow it. She does not trust you and is so terrified that she locks herself inside the bathroom leaving the child outside the bathroom. The only way to persuade her to lend you her car is to twist the child's arm so that she can hear the child scream. Utilitarianism says that you ought to twist the child's arm, since it would be worse to let the injured passenger die.[2]

The judge

A murder has been committed and most people believe that Jake is guilty, but the judge knows he is innocent. If the judge does not get Jake hanged, there will be a riot and several people will die. Utilitarianism says that the judge ought to sentence the innocent Jake to death since causing one to be killed is better than allowing many to be killed.[3]

The transplant

A doctor has five patients who will all die if they do not get an immediate transplant. One patient needs a new heart, two need a new lung, and two need a new kidney. By sheer coincidence, the doctor finds out that a healthy person, who is in hospital for a routine check-up, happens to be the perfect donor for all five patients. Utilitarianism would tell the doctor to cut up the healthy person and distribute the organs to the five patients, since that would maximize total well-being.[4]

UTILITARIAN RESPONSES

A common utilitarian reply is to concede that there are no constraints, but claim that, normally, it is wrong to lie, break promises, kill, and harm the innocent, for these actions will not normally maximize total well-being. It is only in special circumstances that total-well-being is maximized by violating these principles. So, the utilitarian will honour these principles as important rules of thumb rather than constraints on actions. If you follow these principles, you will normally maximize total well-being.

There may even be a good utilitarian argument for adopting a policy of not even thinking about the most repugnant violations as viable options. For instance, it might be better if a judge adopts a policy of not even entertaining the option of punishing the innocent, since such thinking may very well lead to bad consequences. A judge who sees punishing the innocent as 'just another option' will be less sensitive to the rules of law and, since it is difficult for the judge to fake it, this will tend to erode people's trust in law and order. Furthermore, a judge can easily misjudge the situation. Is it certain that the riot can be prevented only by sentencing an innocent person to death? Can he not, for instance, somehow pretend to have him

executed instead? And what about the reactions of the friends and family of the innocent person who is killed?

Similarly, a doctor who is always on the look out for healthy patients to cut up and use as (unwilling) donors has adopted a policy that tends to have very grave consequences in the long run. If people get to know about his policy, they will lose all trust in the doctor and avoid visits to the hospital even when they have an urgent need to be treated. Furthermore, the doctor can easily misjudge the situation. Is it, for instance, so clear that no one will know about his enforced donations? And is it clear that killing one will in fact lead to a full recovery of all his other patients? Aren't there other better policies to consider, for instance, encouraging people to donate their organs after their death so that there will be no need to kill one patient in order to save others?

For less abhorrent violations, such as lying and breaking promises, the utilitarian could claim that it may be better to adopt a policy of thinking of them as actions one 'must not do', at least in situations where no catastrophic outcomes are at stake. This can be compared to the way dieters should think about whether to have another piece of cake. Of course, one more piece of cake will not in itself do any harm, but if you take this to heart and always give yourself the permission to have another piece of cake, it is more likely that you will end up eating the whole cake. It may be better to think 'I must not have another piece of cake' even though you know that another piece of cake would not in fact cause you any harm. Analogously, even if a particular lie would not harm anyone, if you start giving yourself permission to lie whenever you cannot see any harm coming, you will weaken your general commitment to honesty and that will have bad consequences in the long run. Indeed, one might even argue that in order to succesfully participate in the promise institution in the first place you have to block out thinking in terms of small benefits and losses of keeping a promise. If you are constantly calculating the pros and cons of keeping your promise to me, you will no longer be playing the promise game.

One could complain that this utilitarian policy will split the mind of the moral agent, since he will think in ways he knows are false. Indeed, one might worry that it will lead to flat-out inconsistency in his beliefs: he believes that there are no constraints *and* that breaking a promise is simply forbidden. But perhaps the agent does not have

to believe that promise breaking is wrong in order to reap the good consequences of his promise-keeping strategy. The smoker who is trying to kick his habit can combat his urge to smoke by thinking that smoking will certainly kill him when in fact he knows that this is an exaggeration and that there is only a non-negligible probability that he will die from smoking. Perhaps the utilitarian agent can do something similar in his pursuit of total well-being. His strategy is to exaggerate and think 'This is simply impermissible' in cases where he can easily break a promise.

Of course, these replies do nothing to block the conclusion that utilitarianism will sometimes tell you to do repugnant things, for, strictly speaking, there are no constraints and special duties. These replies will only show that the best utilitarian policy is often not to consider the repugnant actions at all or, when you do consider them, to think about them as actions you must not do.

SPECIAL OBLIGATIONS TO OUR NEAREST AND DEAREST

Common-sense morality and many deontologist theories embrace special duties to our near and dear, including duties to our friends, family members (parents, siblings, children . . .), partners (wives, husbands . . .), and fellow members of your community. These duties are grounded in facts about our relations to others. They differ from constraints in that special duties are owed to some people but not to others, depending on which relationship we stand to them. For instance, I owe it to my children to shelter them simply because they are my children. Constraints, however, are owed to anyone (if they are owed at all). I should not torture an innocent person, no matter whether he is a friend, child, parent, or a stranger.

Since special duties are supposed to sometimes override considerations about well-being, the utilitarian cannot accept these duties. For instance, in the choice between saving your child and some stranger's child, you are required to save the stranger's child, if that would produce more total well-being. No special weight is given to the fact that one of the children is your child.

UTILITARIAN RESPONSES

As in the case of constraints, the utilitarian can argue that special duties are good rules of thumb. There are several reasons why it is

often a good utilitarian policy to show more concern for your nearest and dearest. Sidgwick pointed out the following three reasons.[5] First, we tend to derive more pleasures from interactions with our nearest and dearest than from interactions with complete strangers. Taking out a friend for dinner is more enjoyable than taking out a stranger. Second, we often have more knowledge about how to benefit our nearest and dearest than how to benefit strangers, since we know more about our friends' tastes and preferences than those of strangers. Third, we are often in a better situation to distribute benefits to our nearest and dearest. I can more easily visit my ill friend than visit an ill stranger far away.

So, the utilitarian can argue that special duties and utilitarian duties often coincide. However, the coincidence is far from perfect, especially in our modern society where we have both the information and the technological means necessary to help perfect strangers far away. Since many of these strangers are in great distress – they are starving, ill, or very poor – the choice between taking your friend out for dinner and sending the equivalent sum of money to a trustworthy charity should be clear to the utilitarian. Here the demandingness objection to utilitarianism reappears.

There is, however, another way for the utilitarian to make room for special duties. Instead of seeing them as moral duties, he could see them as *non-moral* requirements stemming from certain perfectionist values. In general, something has perfectionist value if it is a good instance of its kind. So, for example, a knife that cuts well is a good instance of its kind and, therefore, a good knife ought to cut well. By the same reasoning, one could say:

a good friend ought to give priority to his friends,
a good parent ought to give priority to his children,
a good partner ought to give priority to his partner, and
a good community member ought to give priority to his community.

Now, if these requirements are normative in the sense that they provide non-moral reason to act, and if these reasons are strong enough to sometimes trump moral ones, utilitarians can happily accept that we sometimes have more overall reason to be a good parent than to do what is morally right, and, therefore, more reason to save our own child rather than a stranger's child.

But this response is convincing only if it is reasonable to reject the thesis that moral reasons are overriding, and, as we noted in the previous chapter, this is controversial. Furthermore, this response still denies that we have *moral* reasons to be partial towards our nearest and dearest, and this is counter-intuitive.

CAN CONSTRAINTS AND SPECIAL OBLIGATIONS BE BUILT INTO THE GOOD?

We have seen that utilitarianism is forced to reject constraints and special duties. At most, it can accept that constraints and special duties have instrumental value. This, no doubt, makes utilitarianism a less intuitive moral theory. But recall that what we have in mind here is *classical* utilitarianism. It is therefore important to ask whether there is another version of utilitarianism that could accommodate constraints and special duties. I shall argue that the answer is both yes and no.

To see why an affirmative answer seems right, consider again the example about the car accident now put in a diagrammatical form in Table 8.1. (For simplicity, I ignore the well-being of you and the woman. We can assume that you and the woman would suffer to the same degree no matter what you did.)

Classical utilitarianism says that the child ought to be tortured since this will maximize the total sum of well-being (−5 compared to −10). But other versions of utilitarianism need not have this implication. For instance, the fact that a person is intentionally harmed can be assigned negative intrinsic value. It is not just bad for a person to suffer; it is also bad for a person that he is intentionally and knowingly harmed. If this more objective conception of well-being is

Table 8.1

Actions	Outcome 1	Total value
Twist the child's arm	The child suffers (−5) Your passenger is saved (0)	−5 + 0 = −5
Not twist the child's arm	The child does not suffer (0) Your passenger is left to suffer (−10)	0 −10 = −10

Table 8.2

Actions	Outcome 1	Outcome 2	Total value
Twist the child's arm	The child suffers (−5) Your passenger is saved (0)	The child is intentionally harmed (−10)	−5 + 0 − 10 = −15
Not twist the child's arm	The child does not suffer (0) Your passenger is left to suffer (−10)	The child is not intentionally harmed (0)	0 −10 + 0 = −10

accepted, utilitarianism can say that we ought to refrain from torturing the child, as Table 8.2 shows.

In this example, not twisting the child's arm will maximize value (−10 compared to −15.) Similar reasoning applies to the other deontological constraints: telling a lie, breaking a promise, and punishing the innocent can all be assigned intrinsic disvalue. Of course, it will sometimes be difficult to provide a plausible story that explains why being violated in some of these ways, being lied to, for instance, must *in itself* be bad for oneself. The non-utilitarian consequentialist, however, need not be worried by this, since he can accept that something is bad without it being bad for anyone. So, even if utilitarians will be hard-pressed to come up with a good justification for all constraints, the non-utilitarian consequentialist may have an easier time providing such a justification.[6]

To see how a non-classical utilitarian can incorporate special obligations to our nearest and dearest, consider again the case of the drowning children. Your child has been playing in the water with a stranger's child and they have now drifted out into deep waters. They are both drowning but, unfortunately, you can only save one. See Table 8.3.

Table 8.3

Options	Outcome
Save your child, let the other child drown	Your child is saved and leads a happy life
Save the other child, let your own child drown	The other child is saved and leads a happy life

Table 8.4

Options	Outcome 1	Outcome 2	Total value
Save your child, let the other child drown	Your child is saved and leads a happy life (10)	Your child is saved by his parent (5)	10 + 5 = 15
Save the other child, let your own child drown	The other child is saved and leads a happy life (10)	The other child is not saved by his parent (0)	10 + 0 = 10

Classical utilitarianism says that both actions are right, if the children's happiness would be the same and the parents' suffering would be the same. But other versions of utilitarianism need not say this. The fact that a child is saved by her own parent can be assigned positive intrinsic value. More generally, it is good that children are saved by their parents because this expresses the intrinsic value of parental care. This is consistent with utilitarianism if we add that it is *good in itself for children* to be the object of parental care. If this more objective conception of well-being is accepted, then even utilitarianism can say that you ought to save your child, as Table 8.4 shows.

Similar reasoning applies to the other special obligations. Again, non-utilitarian consequentialist may have an easier time providing a plausible justification for taking into account the values of special duties, since they do not have to say that these values are necessarily part of someone's well-being.

So far so good, but here is still a sense in which utilitarianism, indeed any form of consequentialism, is unable to incorporate both constraints and special obligations. Defenders of constraints often claim that you are not allowed to violate a constraint in order to prevent other violations of the same constraint. This means, for instance, that you are not allowed to torture one person in order to prevent others from torturing, and you are not allowed to break one promise in order to prevent others from breaking promises. Similarly, defenders of special obligations claim that you are not allowed to violate a special duty in order to prevent other violations of the same special duty. So, you are not allowed to violate your duty to your child in order to prevent others from violating their duties to their children.

This feature of constraints and special duties prevent the utilitarian and, more generally, any kind of consequentialist, from accommodating them. To see this in the case of constraints, consider the schematic example in Table 8.5.

Assume that if I torture A, you will spare B. But if I do not torture A, you will torture B. Now, a torture-sensitive consequentialism could assign intrinsic disvalue to the fact that someone is tortured (and a utilitarian could add that it is in itself bad for someone to be tortured). But this theory will not forbid me to torture A, if the involved torturings are equally bad. For then the outcomes will contain the same good and bad things: that one person is not tortured and that another person is tortured. The outcomes of my options must therefore have the same value and I am thus permitted to torture A. For a consequentialist, it cannot matter that someone is tortured by *me*. What matters is that *someone* is tortured by *someone*. Analogous reasons apply to the other constraints.

The similarly structured example in Table 8.6 shows that no consequentialist can accommodate special duties. Assume that if I save my child, this will prevent you from saving your child, and if I do not save my child, you will save yours. A parental duty-sensitive consequentialist that assigns intrinsic value to the fact that a parent saves his child (and perhaps adds that this is good for the child) will not require me to save my child, since the outcomes of my options

Table 8.5

My options	Outcomes
I torture A	A is tortured, you do not torture B
I do not torture A	A is not tortured, you torture B

Table 8.6

My options	Outcomes
I save my child	My child is saved by me, you do not save your child
I do not save my child	My child is not saved, you save your child

are equal in value. The outcomes contain the same good and bad things: that one child is saved by his parent and that another child is not saved by his parent. For a consequentialist, it cannot matter that *I* save *my* child. What matters is that a child is saved by her parent. Analogous reasons apply to the other special duties.

We now have a clear explanation for why utilitarianism and, more generally, consequentialism, cannot accommodate constraints and special duties. According to all forms of consequentialism, the agent's relation to a violation of a constraint or a special duty does not matter. It does not matter whether *he* is violating the constraint or the special duty. It only matters that these violations would be brought about by his actions. By contrast, according to common sense and many deontologists, the fact that you will violate a constraint or a special duty provides a special reason for you not to violate the constraint or the special duty. This is sometimes summed up in the slogan that consequentialism has no room for *agent-relative* reasons. A constraint-sensitive or special duty-sensitive consequentialism is still agent-neutral in the sense that it does not matter who is doing the killing, the torturing, the promise breaking, or the violation of the special duty. You should simply minimize the number of violations of constraints and special duties, even if that requires that you commit a violation, because violations are intrinsically bad and you should minimize what is bad. The fact that *you* will commit a violation is not morally relevant.

TRADE-OFF PROBLEMS

A constraint-based morality does not allow us to violate constraints whenever that will maximize overall value. But how strict is this constraint supposed to be? Imagine that a lot of suffering is at stake and that the only way to prevent it is to violate a deontological constraint. For example, assume that the only way to avoid the painful end of humanity in a nuclear holocaust is to torture one child. Is it reasonable to say that the constraint against torturing the innocent is sacred even in this extreme situation? This seems too rigid. Even Robert Nozick, a staunch defender of constraints (in fact, he coined the term 'side-constraint'), wavers at this point: 'The question of whether these side-constraints are absolute, or whether they may be violated in order to avoid catastrophic moral horror, and, if the latter, what

the resulting structure might look like, is one I hope largely to avoid.'[7] This elusive attitude might be alright given Nozick's purposes in his book *Anarchy, State, and Utopia*. But the question cannot be avoided when constructing a general moral theory.

Of course, we could avoid this rigid view by qualifying the constraints. They should not just read: do not do x. They should contain exception clauses: do not do x, except in circumstances c1, c2 . . .. But then we need to know how to complete this list. Alternatively, we might say that the constraints are not absolute and allow that given that enough of overall badness is at stake we are allowed to violate the constraint. But then we will have the problem of deciding exactly how much badness must be at stake for us to be allowed to violate the constraints. Where shall we set the threshold?

Deontologists must not just tackle the problem of how to trade constraint violations against the overall good, they also face the problem of how to weigh one constraint violation against another. The deontologist says that we are not allowed to violate one constraint in order to prevent other people from violating constraints. But does he mean that we are not allowed to violate one constraint in order to prevent others from violating *any* kind of constraint? Or does he only mean that we are not allowed to violate one constraint in order to prevent others from violating the *same kind* of constraint?

The first idea is an unreasonably strict theory, for it would not allow me to lie in order to prevent other people from torturing someone, to break a promise in order to prevent someone from killing an innocent person, or to break into someone's house (and thereby damage their property) in order to prevent other people from raping a person. Constraints must therefore be ranked in order of importance, and we should be allowed to violate a less important constraint in order to prevent others from violating constraints that are more important. If this is true, then, of course, the deontologist owes us a justification for treating one constraint as more important than another.

A more important problem is that if this more flexible view is accepted, one can wonder whether even the second idea is acceptable, the idea that I am not allowed to violate a constraint in order to prevent others from violating the *same kind* of constraint. If I am allowed to violate a constraint in order to prevent violations of a slightly less important constraint, why am I not also allowed to violate a constraint in order prevent more violations of the *same* constraint?

CAN CONSTRAINTS AND SPECIAL DUTIES BE JUSTIFIED?

On the face of it, it sounds paradoxical to deny that you can be permitted to commit a violation in order to prevent violations of the same kind. After all, since we think there is something morally undesirable and repugnant about violations, it seems sensible to reduce the number of violations.

The deontologist could reply that this ignores the special role of the agent. It is true that certain acts-types are morally repugnant no matter who performs them, but when a particular agent performs a morally repugnant act, it is *he* who is tainted by the performance of the repugnant act, and thus it is *he* who will bear this moral cost. In contrast, when the agent merely allows other people to do repugnant things, it is they, not he, who will be tainted by doing something repugnant. The agent should therefore avoid the moral cost by refusing to perform a repugnant act even if this means that others will be doing the same kind of repugnant act.

The problem with this reply is that it is too agent-focused: The fundamental reason why an agent should not do something repugnant is that he should not dirty his hands by touching something morally repugnant. As Donald Regan puts it, 'the agent is encouraged to indulge in a sort of Pontius Pilatism, taking the view that as long as he keeps his hands clean, the other agents as well as the consequences can take care of themselves'.[8] But what about the victims who will be allowed to suffer at the hands of other people? Do they have no claim on the agent to be saved from this treatment? To make this objection more vivid, suppose that the only way I can prevent you from being tortured for several days by a group of sadists is by torturing you for a few hours. You are begging me to go ahead and torture you, but I staunchly refuse to do it, since I do not want to be tainted by doing this repugnant act. In this case, I seem to show too much respect for constraints and too little respect for you.

A victim-based deontology would instead say that the reason why I am not allowed to commit a violation in order to prevent other violations is that this is the only way I can respect the true moral status of each person. Each person is *inviolable* in the sense that he cannot be permissibly violated, at least not without his prior consent, in order to prevent similar violations of others. In the example above, I do not impermissibly violate you, because you consented to be tortured by me in order to avoid a greater evil for yourself.

In a case where I can violate you in order to prevent violations of others, it is true that by refraining from violating you I will allow that others are violated. But this does not diminish their moral status, for it is still impermissible to violate them. Moral status defines what we can permissibly do to people, not what actually happens to them. In contrast, if it was permissible for me to violate you in order to prevent violations of others, as utilitarians and consequentialist would have us think, then *everyone's* moral status would be diminished, not just your moral status, for it would now be true that *any* person could be permissibly violated in order to prevent others from being violated. So, on this approach, it is not the fact *I* will do something repugnant and the victim is *mine* that explains why I must not do something repugnant. The explanation is instead that *every* person is inviolable and thus cannot be permissibly violated in order to prevent others from being violated.[9]

This is an intriguing idea, but it should be noted, first, that this kind of inviolability comes in degrees. A person has a maximum degree of inviolability when we are not allowed to violate him no matter how many other violations or how much suffering we could thereby prevent. But such inviolability seems too extreme; it would not allow us to torture one innocent person in order to prevent billions of other people being tortured. Any moderate deontologist who thinks that it is permissible to violate one person when a sufficiently large number of violations or a sufficiently large amount of suffering is at stake would not assign maximum inviolability to people. How much inviolability should we then assign to people? Well, utilitarianism has one answer: no one can be permissibly harmed in order to prevent a lesser amount of harm to others. So, both utilitarians and moderate deontologists assign a less than maximum degree of inviolability to people.

Second, this view simply assumes that moral status is only about what can permissibly be *done* to people. But why isn't moral status also about what can permissibly be *allowed to happen* to people? If you have moral status, how can I be permitted to allow others to violate you? In short, this victim-based approach seems to assume that there is a morally relevant distinction between *doing* and *allowing*. Utilitarians and consequentialists would of course reject such a distinction, since they judge actions by the values of their outcomes, not by the way these outcomes are brought about. In the next chapter, we will take a closer look at the issue whether the way outcomes are brought about makes a moral difference.

CONCLUDING REMARKS

There are no moral 'no go' areas for a utilitarian. Any kind of action, no matter how intuitively repugnant, can be morally right if the choice facing the agent is sufficiently tragic. Also, since everyone counts for one and not for more than one, there is no room for special duties in a utilitarian theory. However, the utilitarian does not deny that *thinking* in terms of constraints and special duties will often have beneficial consequences. So, constraints and special duties can definitively be seen as something instrumentally valuable. They can even be seen as intrinsically valuable if classical utilitarianism is abandoned in favour of a utilitarian theory that adopts a more objective account of well-being (or a non-utilitarian consequentialist theory). But even on this revised utilitarian view, indeed, on any consequentialist view, it is still true that violations of constraints and special duties ought to be minimized, and this may require the agent to 'dirty his hands'. However, it is very difficult to find a plausible explanation of constraints and duties if we maintain that we are never allowed to violate one in order to minimize a great number of similar violations.

SUGGESTED READING

On constraints:

Kamm, F. (2007), *Intricate Ethics*, Oxford: Oxford University Press, 17–23.
Nagel, T. (1986), *The View from Nowhere*, New York: Oxford University Press, 175–180.
Nozick, R. (1974), 'Moral constraints and moral goals', in *Anarchy, State and Utopia*, New York: Basic Books, Chapter 3, 28–30.
Scheffler, S. (1994), *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions,* second revised edition, Oxford: Oxford University Press, esp. Chapters 4–5.

On special obligations:

Jeske, D. (2001), 'Friendship and reasons of intimacy', *Philosophy and Phenomenological Research,* Vol. 63, 329–346.
Scheffler, S. 'Relationships and responsibilities', *Philosophy and Public Affairs,* Vol. 26, 189–209.

On agent-neutral and agent-relative reasons:

Broome, J. (1995), 'Skorupski on agent-neutrality', *Utilitas*, Vol. 7, 315–317.

Nagel, T. (1986), *The View from Nowhere*, New York: Oxford University Press, 152–153, 158–163.
Parfit, D. (1992), *Reasons and Persons*, Oxford: Oxford University Press, 143.

On consequentialism and constraints:

Kagan, S. (1991) 'Replies to My Critics' *Philosophy and Phenomenological Research* 51, 919–920.
Sen, A. (1982), 'Rights and agency', *Philosophy and Public Affairs,* Vol. 11, No. 1, 3–39.
Sosa, D. (1993), 'Consequences for consequentialism', *Mind*, Vol. 102, 101–122.

On consequentialism and special obligations:

Arneson, R. J. (2003), 'Consequentialism vs. special-ties partiality', *The Monist,* Vol. 86, 382–401.
Sen, A. (1982), 'Rights and agency', *Philosophy and Public Affairs,* Vol. 11, No. 1, 3–39.
Sidgwick, H. (1907), *The Methods of Ethics*, seventh edition, London: Macmillan, Book 4, Chapter 3.