Beyond Action:

Applying consequentialism to decision making and motivation

Toby Ord
Balliol College & Christ Church
University of Oxford
March 2009

Thesis for the degree of Doctor of Philosophy

Abstract

It is often said that there are three great traditions of normative ethics: consequentialism, deontology, and virtue ethics. Each is based around a compelling intuition about the nature of ethics: that what is ultimately important is that we produce the best possible outcome, that ethics is a system of rules which govern our behaviour, and that ethics is about living a life that instantiates the virtues, such as honesty, compassion and loyalty. This essay is about how best to interpret consequentialism. I show that if we take consequentialism beyond the assessment of acts, using a consequentialist criterion to assess decision making, motivation, and character, then the resulting theory can also capture many of the intuitions about systems of moral rules and excellences of character that lead people to deontology and virtue ethics.

I begin by considering the argument that consequentialism is self-defeating because its adoption would produce bad outcomes. I take up the response offered by the classical utilitarians: when properly construed, consequentialism does not require us to make our decisions by a form of naïve calculation, or to be motivated purely by universal benevolence. Instead it requires us to use the decision procedure that will produce the best outcome and to have the motives that lead to the best outcome. I take this idea as my starting point, and spend the thesis developing it and considering its implications.

I demonstrate that neither act-consequentialism nor rule-consequentialism has the resources to adequately assess decision making and motivation. I therefore turn to the idea of global consequentialism, which assesses *everything* in terms of its consequences. I then spend the greater part of the essay exploring how best to set up such a theory and how best to apply it to decision making and motivation. I overcome some important objections to the approach, and conclude by showing how the resulting approach to consequentialism helps to bridge the divide between the three traditions.

Acknowledgements

There are many people who have contributed a great deal to this thesis and to whom I owe considerable gratitude.

I would first like to acknowledge the two colleges which have provided my financial support, my home, and much more during my time at Oxford. Balliol College supported me for my first three years at Oxford: during the BPhil (where the foundations for the present thesis were laid) and the first year of my DPhil. Christ Church has then supported me for the past two and a half years, with a very generous scholarship. In their own distinctive ways, they have each provided me with a wonderful environment in which to live and study.

This thesis was supervised by John Broome and Derek Parfit. I could not wish for better mentors. They have provided me with a wealth of insightful comments on my thesis and my other projects, and it has been both a pleasure and an honour to work with them.

There are many other people with whom I have discussed this work and whose insights have found their way into the finished thesis. I would particularly like to thank: Gustaf Arrhenius, Nick Bostrom, Krister Bykvist, Roger Crisp, Tom Douglas, Brad Hooker, and Pablo Stafforini; as well as audiences from the Department of Practical Philosophy at Stockholm University, and the Tenth Conference of the International Society of Utilitarian Studies at Berkeley.

Finally, I wish to thank my wife, Bernadette Young, for her constant support and encouragement. In a project of this scale, one needs a source of stability and a source of inspiration. She was both.

Contents

1 Introduction **1**
1.1 Act-consequentialism 1
1.2 Could consequentialism be self-defeating? 4
1.2.1 The 'paradox' of hedonism 5
1.2.2 The 'paradox' of benevolence 7
1.3 The standard consequentialist reply 9
1.4 The structure of this essay 13

2 Do the traditional forms of consequentialism suffice? **16**
2.1 Act-consequentialism and decision procedures 16
2.1.1 Decisions as acts 17
2.1.2 Implicit obligations 19
2.1.3 Evaluation of decision procedures 22
2.1.4 The act of adopting a decision procedure 22
2.2 Rule-consequentialism 24

3 Applying consequentialism to everything **29**
3.1 Evaluative focal points 29
3.2 Four rough accounts of global consequentialism 32
3.2.1 The four accounts 32
3.2.2 Focal points and roles in which to evaluate them 34
3.2.3 Right or best? 37
3.2.4 Formalizing role-based global consequentialism 38
3.3 Proposition-based global consequentialism 42
3.3.1 Possibilist global consequentialism 42
3.3.2 An actualist account 48
3.4 Strengths and weakness of global consequentialism 50
3.4.1 The merits of global consequentialism 50
3.4.2 Objections to global consequentialism 52

4 Interpreting the question of decision making **55**
4.1 Prior approaches 56
4.1.1 Habits and predispositions 56
4.1.2 Rules and principles within decision making 58
4.2 The nature of decision procedures 61
4.3 'Following' a decision procedure 64
4.3.1 Compliance 64

4.3.2 Perfect execution 66
4.3.3 Commitment 68
4.4 Further clarifications 70
4.4.1 Which type of rightness 70
4.4.2 Scope 72
4.4.3 For one or for all? 73

5 Motivation and Character 76
5.1 Do the traditional forms of consequentialism suffice? 76
5.1.1 Act-consequentialism 76
5.1.2 Motive-consequentialism 78
5.2 Motives and character 79
5.3 Roles 84
5.4 Further clarifications 88

6 Problems addressed 91
6.1 The inconsistency objection 91
6.1.1 Adams and Hooker 92
6.1.2 Streumer 94
6.2 A regress concerning decision procedures 96
6.3 The indistinctiveness objection 99
6.4 Blame and guilt 102

7 Conclusions 105

Appendix 109
Right or Best? 109

References 117

Chapter 1

Introduction

A common criticism of consequentialism is that if we attempt to apply it in practice, this may predictably lead to worse outcomes than if we were to adhere to some other moral theory. In this chapter, I shall introduce consequentialism, then provide a sketch of the criticism and of the standard consequentialist reply.

This standard reply contains an insight that has the potential to reshape our understanding of consequentialism and its relationship to the other major ethical traditions of deontology and virtue ethics. In the chapters that follow, I shall develop this insight to its fullest extent, producing a robust theory of consequentialism that can be used to assess not just acts, but decision making and motivation.

1.1 Act-consequentialism

Act-consequentialism is, at heart, a very simple account of the connection between the rightness of acts and the goodness of their outcomes. It rose to prominence in the work of the early utilitarians[1] and has had an enormous impact on moral philosophy in the 19th and 20th centuries. During this time, it has been continuously refined by its supporters and has served as a measuring stick to which rival moral theories are compared. Act-consequentialism is, roughly speaking, the following doctrine:

> *The act-consequentialist criterion of rightness*
>
> An act is right *iff* it will lead to at least as much good as any alternative act.[2]

This principle must then be supplemented by an *axiology*: an account of what it is that constitutes the good. For example, the earliest and most well known act-consequentialist theory, *hedonistic utilitarianism*, holds that the good is the total balance of happiness over

[1] Notably: Jeremy Bentham, John Stuart Mill, Henry Sidgwick and G. E. Moore.

[2] The definition here is formulated so as to explicitly take ties into account; however it is often convenient to speak as if there is but a single best act and, since nothing important shall turn on this, I take this convenience herein. I use the term 'right' to mean the same as 'permissible'; however since I am discussing consequentialism and ignoring ties, everything which is permissible is also obligatory. The term 'iff' is to be read as 'if and only if'.

suffering. Other theories claim that the good is constituted in some different way, such as by the fulfilment of rational desires or the possession of certain attributes from some objective list (*e.g.* happiness, education, being loved by someone). Act-consequentialist theories can move further still from the classical utilitarian versions, taking account of the distribution of personal good over the population or, to take an extreme example, ranking the outcomes by the number of broken promises they contain. While the choice of axiology is thus of considerable importance, it will not play a part in this essay. The issues that I wish to discuss are independent of the axiology chosen, and a resolution will be of use to all act-consequentialist theories.

Even ignoring the question of axiology, one finds considerable debate about how the criterion of rightness should be interpreted.[3] For example, some act-consequentialists are concerned with the *causal consequences* of an act, looking only at the good that is caused by the act under consideration. Others interpret the criterion as being concerned with the future that would come about were the act performed, regardless of causal connections. On a third interpretation, the proper objects of evaluation are not mere futures, but entire possible worlds which include a past, present and future. This would let us take past events into account when judging actions, and allow concepts like desert or promise breaking to enter the axiology. On such points I shall try to remain neutral, as I hope my arguments will apply in all interpretations. In what follows I will use the terms 'consequences' and 'outcomes' interchangeably, and neither is to imply anything about whether there is a causal relation involved.

However, there are two major distinctions that I cannot avoid discussing. I have introduced the criterion as one that assesses the rightness of individual acts and does so on the basis of their consequences in the situation at hand. It is thus known as *act-consequentialism*. Opposed to this are various forms of *indirect consequentialism*, where the acts are assessed on the basis of something else. The classic example is *rule-consequentialism* in which rules are assessed on the basis of their consequences and the right acts are those which are in accord with the best set of rules.

We shall begin by considering act-consequentialism, since this is the most common interpretation and the major target of the criticisms that I wish to discuss. Later, we shall consider whether a move to rule-consequentialism or some other form of consequentialism can help. Obviously act-consequentialists will be reluctant to abandon the key aspects of consequentialism that led them to it in the first place. It is thus worth mentioning a key intuition behind act-consequentialism, one that I shall be guided by during the entire discussion to come: that morality is fundamentally about promoting good outcomes. In his work *Reasons and Persons*, Derek Parfit expresses this as follows:

> There is one ultimate moral aim: that outcomes be as good as possible.[4]

---

[3] See Carlson (1995) for a detailed overview.

[4] Parfit (1984) p. 24. This is not to say that people should consciously aim at producing good outcomes, but that producing good outcomes is in some sense the purpose of morality.

Let us call this claim the *fundamental consequentialist intuition*. Later, we shall see that act-consequentialism arguably lacks sufficient resources to reply to its critics, but that there is a simple and intuitive extension called global consequentialism which *can* mount such a reply. Moreover, it can do so without deviating from this fundamental consequentialist intuition. Until I return to the topic of different forms of consequentialism, I shall simply use the term 'consequentialism' to refer to 'act-consequentialism'.

Finally, there is the issue of objective and subjective rightness. The criterion of rightness that I have given is stated in objective terms: an act is right if and only if it leads to the most good. It doesn't matter whether the agent believed that it would lead to the most good, or whether the agent had reason to believe this. All that matters is that the act actually does lead to the most good. This is a useful sense of rightness and it is involved in many judgements about what was right in light of all the information.

However, there are also a range of subjective senses of rightness. For example, we sometimes consider which act would be right in a sense that takes into account the agent's epistemic situation. We can do this by replacing the actual goodness that each act would lead to with the *expected goodness*.[5] Indeed, there are several ways in which we can do this, each leading to its own sense of rightness. We could just consider the beliefs the agent actually has, in which case we get:

> An act is right in the belief-relative sense *iff* it will lead to at least as much expected good as any alternative act, where the probabilities are given by the agent's degrees of belief

Or we could consider the decision in the light of the evidence available to the agent.[6] This will come apart from the former sense for agents who haven't taken full account of the evidence when forming their beliefs:

> An act is right in the evidence-relative sense *iff* it will lead to at least as much expected good as any alternative act, where the probabilities are given by the agent's evidence

The subjective and objective senses of rightness have often been seen as competitors for the

---

[5] This is a basic term from probability theory that refers to the sum of the good produced in each possible outcome, weighted by the probability of these outcomes. For example, if there is a 30% chance that the act will lead to 100 units of good and a 70% chance that it will lead to 10 units of good, then the expected good is 37 units. Note that no-one has to believe or 'expect' that the outcome will actually involve 37 units of good (in this case it can only involve either 10 or 100 units).

Note also that my use of expected good to evaluate uncertain outcomes is implicitly *risk-neutral*. Some people argue that we should instead be *risk-averse* and thus use a different calculation which gives more weight to bad outcomes. I would like to remain uncommitted on this issue, and I only specify 'expected good' for concreteness. Friends of risk aversion can substitute their preferred method of evaluating prospects.

[6] Note that there are several ways in which we could interpret 'evidence available to the agent' and these lead to their own sub-senses of rightness.

position of being rightness *simpliciter*. However, there has been a movement towards treating them as complementary conceptions of rightness: sometimes we are interested in the objectively right act, sometimes in the right act in the belief-relative sense, sometimes in the right act in the evidence-relative sense, and sometimes even in the act that the agent believed was right.[7] By keeping these senses side by side, we can gain a great deal of expressive power and lose nothing of significance.[8] I shall thus adopt this approach in the present essay, using 'right' without any modifier to mean 'objectively right', and inserting the appropriate modifier for any of the subjective senses of right.[9]

1.2 Could consequentialism be self-defeating?

Suppose everyone were to adopt consequentialism. It is often argued that this would have bad consequences. For example, since people are often biased towards themselves, they would tend to underestimate the social costs of lying, promise-breaking, or theft, leading to more of these problematic behaviours than if they were to accept a more rigid moral code. If true, then it would appear that consequentialism is in some sense self-defeating, as its fundamental aim (that things go as well as possible) is worse achieved when we adopt it than when we do not.

Let us call this the *self-defeatingness objection*.[10] There are several different versions of this objection which can be separated by making three distinctions.

The first distinction is whether the adoption of consequentialism is supposed to be worse than the status quo, or whether it is merely supposed to be worse than the adoption of some other moral theory. The latter is a more modest complaint, but if it works, it is sufficient to cause problems for consequentialists. I shall therefore assume this interpretation in what follows.

The second distinction is whether we are considering the adoption of consequentialism by an individual or by everyone. The objection is more frequently cast in terms of the latter, but the main points would work equally well when considering individual adoption of consequentialism and this would be even more damaging to consequentialism if the argument can be made to work. I shall therefore concentrate on this interpretation.

---

[7] See Parfit (1984), pp. 24–40, Gibbard (1990) pp. 42–3, Zimmerman (1996), pp. 10–20.

[8] Readers who believe there can be only one true sense of 'right' and 'ought' can simply focus on what I say about the sense they hold to be central.

[9] Some readers will notice that I have not explicitly dealt with 'objective probabilities' such as those that come up in questions of radioactive decay and (perhaps) in coin tossing. These would fall somewhere between the objective and subjective realms mentioned above, for they would require expected goodness and yet the probabilities would not be relativised to a given set of beliefs or evidence. I have omitted them merely for ease of exposition and to focus on what I see as the more important senses of rightness.

[10] Note that in Parfit's terms this objection is that consequentialism is *indirectly self-defeating* rather than that it is *directly self-defeating*, for it asserts that the aims of consequentialism are worse achieved by its being adopted, not that they are worse achieved by its being successfully followed [Parfit (1984), pp. 5, 53].

The third distinction is more subtle than the others. It concerns what it means for someone to 'adopt' consequentialism. For example, this might mean that they are motivated to produce outcomes that are as good as possible.[11] If so, we could say that they are motivated by *impartial benevolence*. Alternatively, it might mean that they make their decisions by the direct application of the criterion of (belief-relative) rightness. In other words, they would consider possible futures that their acts could lead to, estimate the relative likelihoods of each possibility, calculate the expected good of each available act, and then choose the one with the greatest expectation. Let us call this the method of *naïve calculation*.[12]

These two ways of understanding the adoption of consequentialism turn out to be quite similar in practice, for if motivated to produce the best outcome in any situation, many people would use some form of naïve calculation (perhaps a very rough and ready form), and anyone using naïve calculation to choose their actions is in some sense motivated to produce as much good as possible. However, this distinction will turn out to be very relevant when we come to look at the traditional consequentialist replies to this objection.

First, though, let us explore the numerous ways in which the adoption of consequentialism could predictably lead to bad results: both in terms of one's own happiness and in terms of the overall good.

1.2.1 The 'paradox' of hedonism

We all, to greater and lesser extents, desire happiness and seek it out. However, aiming to increase one's happiness can sometimes predictably lead to less happiness than one would have had in the absence of such an aim. In his autobiography, Mill describes this hard-won lesson:

> 'I never, indeed, wavered in the conviction that happiness is the test of all rules of conduct, and the end of life. But I now thought that this end was only to be attained by not making it the direct end. Those only are happy (I thought) who have their minds fixed on some object other than their own happiness; on the happiness of others, on the improvement of mankind, even on some art or pursuit, followed not as a means, but as itself an ideal end. Aiming thus at something else, they find happiness by the way.'[13]

---

[11] It would be even more accurate to say that 'they are motivated to produce outcomes which have as much *X* as possible', where *X* is a non-evaluative description of the good. In other words, if the good is happiness, that they are motivated to produce outcomes with the most happiness. The same goes for our understanding of naïve calculation.

[12] A third interpretation of 'adopting consequentialism', which I shall not look at in detail, is 'believing consequentialism is true'. This would lead to a weak version of the paradox of benevolence (since some who believe this will no doubt attempt a form of naïve calculation), but would succumb to the same response as the other interpretations in section 1.3: consequentialists have never said that their theory demands that we believe in it, and Sidgwick has very famously denied this [Sidgwick (1907), p. 490].

[13] Mill (1873), Ch. 5.

Henry Sidgwick termed this the *paradox of hedonism*:[14]

> 'Here comes into view what we may call the fundamental paradox of Hedonism, that the impulse towards pleasure, if too predominant, defeats its own aim.'[15]

> '…it seems true that Happiness is likely to be better attained if the extent to which we set ourselves consciously to aim at it be carefully restricted.'[16]

This phenomenon can be clearly seen in many cases and in varying strengths. In particular, we shall see that the stronger of these cases show that naïve calculation — as applied to maximising one's own happiness or pleasure — can be self-defeating. That is, we can see that other ways of making our decisions are superior. Unsurprisingly, this will have significant implications for naïve calculation as applied to maximising the overall good.

The paradox of hedonism typically comes into play when the maximisation of happiness requires one's mind to be in a state that is incompatible with calculation. For example, the enjoyment gained from peaceful relaxation, meditation or quiet reflection is incompatible with active calculation. If we are always calculating how to achieve the most pleasure, then the pleasures inherent in such tasks are unavailable to us. Similarly, if we are to watch a film or read a novel then we mustn't be constantly thinking about whether we are enjoying it (or whether it would be best to stop and do something else), for we will never get sufficiently immersed to enjoy it. The same is true for spending time with a friend or going to a party: questioning the quality of our enjoyment diminishes it.[17]

There are also cases where we need to be in a state of mind that is incompatible with calculation, but where this state of mind is not itself pleasurable. For example, we may be driving a car in difficult circumstances or trying to remember our lines in a performance and if we were also calculating the relative benefits of other courses of action (stopping the car, improvising some lines) then we may quite predictably fail at the task at hand. Even if this does not lead to a loss of pleasure at the time, it may well prevent much pleasure later when we have to pay to repair the car or do not get cast in the next play. There are a great many cases like this, most notably that of falling in love. If we are always weighing up faults and virtues, calculating whether or not we should pursue a relationship, then it is unlikely that we shall ever fall deeply in love and experience all the happiness that this brings.

A somewhat different case is the so-called 'tyranny of choice'. When going to the supermarket, we are frequently confronted with a multitude of choices for every trivial option. Which type of toothpaste should I buy? There are many types of toothpaste which each claim their own different advantage. If we are to calculate the best option from the

---

[14] We shall take this phrase to refer both to the pursuit of pleasure and to the pursuit of happiness, for the problem arises in both cases.

[15] Sidgwick (1907), p. 48.

[16] Sidgwick (1907), p. 405.

[17] See Pettit and Brennan (1986) for a detailed discussion of this and many other cases in which calculation is incompatible with pleasure or other things we might judge to be good.

available evidence in all such cases, then we will spend our lives paralysed by such trivial choices. Much better in such a situation would be to *satisfice*: to accept the first choice which is sufficiently good and then move on.

Related to this are those cases in which the time available to choose is very small. If I am heading out to meet friends and see the bus pulling into the stop while I am still some distance away then I have two options: to run for the bus or to keep walking and catch the next one. However, if I spend the time to calculate which is best then the bus will leave and I will deny myself the possibility of running to catch it.

Finally, consider the following case:

> *The race*
>
> Thomas has enrolled in a long distance footrace and expects to do quite well. His enthusiasm helps him through the final stages and he has a very enjoyable run, coming third overall. As it happens, if he were to have carefully considered his knowledge of his competitors he would have realised that they all had better form than he did and he would have become discouraged. This would have made him unable to keep running so strongly in the final stages and he would not have done very well.

If Thomas had been performing the calculations as to whether running in the race would maximise his pleasure, he would have had to consider his prospects of finishing well. However, this would have led to a poor performance and a decrease in pleasure. In the case of competitions, this effect is often present and quite predictable in nature. Having optimism untarnished by reasoned prediction can frequently lead to a better performance and, at the very least, a more pleasant experience.[18]

1.2.2 The 'paradox' of benevolence

While the cases above have been framed in terms of happiness or pleasure, the underlying principle can easily be seen to affect other quantities that people might try to maximise. We have seen how naïve calculation can lead to lost races, missed appointments, wasted time, car accidents, unappreciated novels and failed relationships. In short, it can lead to a reduction of one's own well-being on any measure we might consider. Furthermore, since well-being is a part of the total good on any plausible axiology, we have also seen how naïve calculation can lead to a reduction in the good.

The above examples were all focused on people aiming at their own happiness and consequently lowering it (thus lowering the total good as well). Unsurprisingly, there are also examples where direct attempts to increase the total good can predictably lead to worse outcomes. I shall refer to the existence of such examples as the 'paradox' of benevolence.[19]

---

[18] In other cases the reverse is true: we predictably do best to dampen our expectations so that disappointment will be lessened and excitement raised.

[19] The examples of the previous section can also be easily modified into true examples of the paradox of

A frequently cited example concerns the special relationships between lovers or friends. Since a great deal of what matters in our lives is generated through such relationships, this is very important indeed. Pettit and Brennan put it thus:

> '…an uncomplicated illustration is provided by the security which lovers or friends produce in one another by being guided, and being seen to be guided, by maxims of virtually unconditional fidelity. Adherence to such maxims is justified by this prized effect, since any retreat from it will undermine the effect, being inevitably detectable within a close relationship. This is so whether the retreat takes the form of intruding calculation or calculative monitoring. The point scarcely needs emphasis.'[20]

Michael Stocker has also developed this point, taking it to be a strong criticism of consequentialism (and many other moral theories):

> 'Love, friendship, affection, fellow feeling, and community, like many other states and activities, essentially contain certain motives and essentially preclude certain others; among those precluded we find motives comprising the justifications, the goals, the goods of those ethical theories most prominent today. … to the extent that you live the theory directly, to that extent you will fail to achieve its goods.'[21]

Another prominent example concerns what Bernard Williams calls *personal integrity*.[22] We each take our own personal projects very seriously and they can often become central to our lives. If we were motivated by impartial benevolence and forced to see other people's projects as just as important as our own, then we would not be able to engage with them in as fulfilling a way. Though Williams may not have known it, this is a development of a more general idea noted by Sidgwick almost a century earlier:

> '…the fullest development of happy life for each individual seems to require that he should have other external objects of interest besides the happiness of other conscious beings.'[23]

Thus to the extent that the adoption of consequentialism requires such an impartial view, it defeats its aims.

A major problem for benevolent calculation concerns the cost of deliberation. If someone is drowning in the river and we pause to calculate, it will increase the chance that they die. The only way to achieve the best outcome in such a case is through some easily applied rule, such as those found in common-sense morality. While this example is extreme, there are a great many cases in which the costs of the time spent calculating predictably outweigh the benefits.

---

benevolence, because people may try to increase the overall good by trying to increase their own well-being and, by undermining this goal, ultimately undermine the overall good.

[20] Pettit and Brennan (1986), p. 450. See also Smart (1973), pp. 44–5 and Railton (1984).

[21] Stocker (1976), p. 461.

[22] Williams (1973), pp. 108–18.

[23] Sidgwick (1907), p. 405. See also Parfit (1984), pp. 27–8.

This may be either when the costs of calculation are extreme, as in the case of the drowning person, or when the benefits of calculation are very low, as in the case of toothpaste selection.

Furthermore, natural biases will affect benevolent calculation even more than in the case of pleasure maximisation. For while we are biased towards our near future over our further future, we tend to be even more biased towards ourselves over others. Practitioners of naïve calculation are unfettered by the standard moral prohibitions and so there are many opportunities for their own biases to lead them to steal or commit other crimes.[24] This problem would be even more pronounced if everyone, or the great majority of society, practiced naïve calculation. While it is not clear that the resulting situation would be worse than our present one,[25] it takes little imagination to suggest decision procedures which improve upon naïve calculation by curtailing its scope and thus avoiding cases where bias is prevalent.

1.3 The standard consequentialist reply

As we have seen, there are several ways in which the adoption of consequentialism could predictably lead to bad outcomes. On the face of it, this seems like a very damaging criticism, for if consequentialism makes demands of people that predictably lead to worse outcomes, then it would fail on its own terms. However, consequentialism does not make such demands. The bad outcomes discussed above are all produced by restricting people's motives to those of impartial benevolence and by restricting their decision making to the process of naïve calculation. For nearly two centuries now, consequentialists have denied that their theory demands either of these things. For example, as early as 1832, John Austin wrote:

> 'It was never contended or conceived by a sound, orthodox utilitarian, that the lover should kiss his mistress with an eye to the common weal.'[26]

Mill explicitly denied that consequentialism required naïve calculation.[27] For example, in his 1838 essay, 'Bentham', he wrote:

> 'Those who adopt utility as a standard can seldom apply it truly except through the secondary principles. … It is when two or more of the secondary principles conflict, that a direct appeal to some first principle becomes necessary; and then commences the practical importance of the utilitarian controversy; which is, in other respects, a question of arrangement and logical subordination rather than of practice; important principally in a purely scientific point of view, for the sake of the systematic unity and

[24] See, for example, Moore (1903), p. 162 and Parfit (1984), p. 28.

[25] Smart (1956), p. 348, points out that the benefits of having the world's governments (and their nuclear weapons) run by people devoted to benevolence would likely overwhelm the disadvantages. To this I would also add the good that would be done by the resulting massive increase in aid to the world's poor.

[26] Austin (1832), p. 108.

[27] The most famous examples are in the second chapter of *Utilitarianism* [Mill (1861)], but those quoted here make the points more clearly.

> coherency of ethical philosophy.'[28]

It is thus clear that he saw the greatest happiness principle as an ultimate unifying principle of ethics, rather than as something to which one must constantly appeal in decision making. This latter role was to be filled by so-called 'secondary principles', such as those of honesty or charity. Furthermore, in his *System of Logic* of 1843, Mill wrote a short section on utilitarianism which was almost entirely devoted to stating his view that utilitarianism did not require us to be motivated by impartial benevolence:

> 'I fully admit that this is true: that the cultivation of an ideal nobleness of will and conduct, should be to individual human beings an end, to which the specific pursuit either of their own happiness or of that of others (except so far as included in that idea) should, in any case of conflict, give way. But I hold that the very question, what constitutes this elevation of character, is itself to be decided by a reference to happiness as the standard. The character itself should be, to the individual, a paramount end, simply because the existence of this ideal nobleness of character, or of a near approach to it, in any abundance, would go further than all things else towards making human life happy…'[29]

In *The Methods of Ethics*, Sidgwick makes a similar point:

> '…the doctrine that Universal Happiness is the ultimate*standard* must not be understood to imply that Universal Benevolence is the only right or always best *motive* of action. For, as we have before observed, it is not necessary that the end which gives the criterion of rightness should always be the end at which we consciously aim: and if experience shows that the general happiness will be more satisfactorily attained if men frequently act from other motives than pure universal philanthropy, it is obvious that these other motives are reasonably to be preferred on Utilitarian principles.'[30]

Indeed, Sidgwick spent a large part of *The Methods of Ethics* discussing the questions of which set of motives and which code of ethical decision making would lead to the most happiness. He evidently saw these as important empirical questions for consequentialists to answer.

This view has continued to find acceptance with consequentialists for more than a hundred years.[31] Indeed, I know of no prominent consequentialist since Bentham who has claimed that consequentialism requires us to be motivated by impartial benevolence or to make our decisions through naïve calculation, and even Bentham's support could be questioned.[32]

---

[28] Mill (1838), p. 111.

[29] Mill (1843), Bk. VI, Ch. 12, §7.

[30] Sidgwick (1907), p. 413. The emphasis is his.

[31] See, for instance: Moore (1903), p. 162–4; Smart (1973), § 7; Bales (1971); Hare (1981), Ch. 4; Parfit (1984) pp. 24–9, 31–43; Railton (1984), pp. 165–8; Pettit and Brennan (1986); Driver (2001). I shall address their individual approaches in Chapters 4 and 5.

[32] See for example his comments immediately following his felicific calculus: 'It is not to be expected that this process should be strictly pursued previously to every moral judgement, or to every legislative or

The best-known modern account is given in the paper, 'Act-utilitarianism: account of right-making characteristics or decision making procedure?' by Eugene Bales.[33] In it, Bales takes great care to distinguish between a *criterion of rightness*, which specifies whether an act is right, and a *decision procedure*, which is a practical method for choosing which act to perform. He notes that a criterion of rightness need not specify the right act in terms that enable us to pick it out. For example, it could specify that the right act is the act that God commands even if we cannot tell which act this is. Bales treats a criterion of rightness as physicists treat the fundamental physical laws: these laws govern the facts of the world (which acts are right, or where the planets are located at a given time), but it may not be expedient for us to use them to calculate these answers.[34]

A decision procedure, on the other hand, is a practically applicable method for deciding what to do. It may involve appeals to principles or rules, but only so far as they are practically applicable. A decision procedure may take the form of a precise sequence of steps or it may take the form of an intuitive process that is difficult to articulate. Either way, it must be capable of guiding our actions. Later on I shall spend considerable time exploring the nature of decision procedures, but for now let us leave it as an intuitive concept.

Bales argues for a form of utilitarianism in which the greatest happiness principle serves only as a criterion of rightness and in which the appropriate decision procedure is whatever decision procedure would lead to the most happiness. If the decision procedure of naïve calculation leads to less overall happiness than some other decision procedure, then we should not use naïve calculation. Since consequentialists are not required to use naïve calculation, the objection that it would be self-defeating to do so becomes irrelevant.

This distinction between criteria of rightness and decision procedures is now the widely accepted way for consequentialists to look at this issue. However, there is something lacking in this analysis. It shows how consequentialism does not require the use of naïve calculation, but says nothing about impartial benevolence. Instead of a two-fold distinction between criteria of rightness and decision procedures, we really need a three-fold distinction between criteria of rightness, decision procedures *and patterns of motivation*. This third aspect arose in the challenge that consequentialism is self-defeating and, as we have seen above, it was addressed by Mill and Sidgwick in much the same way as they (and Bales) addressed decision procedures: the appropriate set of motives to have is the set of motives that will lead to the best outcome.[35]

---

judicial operation.' [Bentham (1789), Ch. IV, par. VI].

[33] Bales (1971).

[34] cf. Mill's famous comments on the Nautical Almanack [Mill (1861), Ch. 2].

[35] This consequentialist approach to motivation has also famously been put forward by Robert Adams in his influential paper 'Motive utilitarianism' [Adams (1976)]. However, Adams did not address the question of decision making and thought that motive-consequentialism was ultimately incompatible with act-consequentialism.

Let us now bring all of this together by distinguishing between three kinds of question that a moral theory might answer:[36]

1. *The question of action:* 'What should I do?'[37]
2. *The question of decision making:* 'How should I decide what to do?'
3. *The question of motivation:* 'How should I be motivated?'

The (act-)consequentialist criterion of rightness is a reply to the question of action, determining the act an agent should perform in any particular circumstance. The question of decision making, however, does not judge acts themselves. Instead, it judges the individual ways in which an agent can make her decisions. It is typically assumed by deontologists that the answer to the second question is intimately connected to the first. For example, that the obligatory actions are those that do not break any of a certain set of rules and that we should decide what to do by testing our actions against these rules.

However, for consequentialists it seems much more promising to answer both of these questions through an appeal to the fundamental consequentialist intuition: I should do whichever act leads to the most good and I should decide what to do in whichever way leads to the most good. In other words, I should follow whichever decision procedure it is that leads to the most good. Perhaps this decision procedure involves following a certain set of rules, or reflecting upon the moral virtues, or considering whether the maxim behind a proposed act can be universally willed. Perhaps the best decision procedure is a small modification of common-sense morality, or a method that no moral theorist has yet discussed. For a consequentialist this is an open empirical question.

The third question concerns patterns of motivation: psychological features involving desires and dispositions, that cause agents to act in certain ways. I shall also take this question to be concerned with a person's character, which is closely related to motivation though arguably distinct. Virtue ethicists hold that the question of motivation (or something like it) is central to understanding ethics: that the correct focus is on the motivation or character of the agent. Certain character traits are known as virtues, such as honesty, compassion, or integrity; and the focus of ethics lies in identifying such virtues and instilling them within ourselves, leading us to attain the ideal character.[38]

However, as above, consequentialists will find it promising to answer this question in the same way as the other two: by an appeal to the fundamental consequentialist intuition. I

---

[36] Adams [(1976), p. 474] makes a similar distinction, but with only two questions. He uses exactly the same phrasing for the question of action ('What should I do?') and similar phrasing for the question of motivation ('What motives should I have?').

[37] Here, and throughout the essay, I am using the term 'should' synonymously with 'ought', expressing a moral obligation. I use 'should' rather than 'ought' merely because it sounds more natural.

[38] This explanation of motivation, character, virtue, and virtue ethics is of necessity very brief and thus does not do justice to the subtleties involved. Much more detail can be found in Chapters 5 and 7.

should be motivated in whichever way it is that will lead to the most good. It is highly plausible that this will involve patterns of motivation or character traits corresponding to many of the widely held virtues, such as honesty and compassion. It is also likely that it will involve agent-centred motives — such as love and friendship — alongside a degree of universal benevolence. However, it must be stressed that for the consequentialist the question of which are the ideal motives or character traits is an empirical one, dependent entirely on whichever motives or traits will produce the most good.

This essay is an attempt to flesh out this consequentialist approach to motives and decision procedures. In the consequentialist literature, much emphasis is placed on the question of action at the expense of the questions of decision making and motivation. In this essay, I take a modest step towards rectifying this, showing in some detail how the historical consequentialist approach to these issues can be developed and made precise, exploring the challenges that beset it and the new light it casts upon the nature of consequentialism. In particular, I aim to show that this approach can partially reconcile consequentialism with both deontology and virtue ethics, explaining away much of their apparent disagreement.

In the formulations above, the questions of decision making and motivation have been left somewhat vague. This is a deliberate choice. In Chapters 4 and 5, I shall spend considerable time clarifying the concepts involved, but for now it is enough to see that there *are* important questions for a moral theory to answer concerning decision making and motivation.

1.4 The structure of this essay

This essay examines the consequentialist approach to answering the question of decision making and the question of motivation. Since there is much overlap in the approach to these questions, I shall take them one at a time. Chapters 2–4 will thus take us through the process of determining how we should decide what to do. Chapter 5 will then transfer these results to the question of how we should be motivated. Chapter 6 will address objections to this approach, and Chapter 7 will examine what it means for the relationship between consequentialism, deontology and virtue ethics.

Chapter 2

I look at ways in which the traditional forms of consequentialism might hope to talk of the rightness of decision procedures alongside rightness of acts. I will first argue that act-consequentialism does not have sufficient resources to judge the rightness of decision procedures as well as the rightness of acts. In effect, this means that the question of decision making is beyond the scope of act-consequentialism. I will then consider its well-known rival, rule-consequentialism, and conclude that this too fails to provide the necessary account.

Chapter 3

I introduce an approach known as global consequentialism, in which acts, rules, decision procedures, motives (and all other focal points) are assessed directly in terms of the good that

they lead to. I expand the theory to take into account the *role* in which each focal point is to be assessed, and then demonstrate how this expanded theory can incorporate a promising account of rightness for decision procedures, while upholding the fundamental consequentialist intuition. This is by necessity a lengthy and rather technical chapter, as it provides the theoretical underpinnings of my project.

Chapter 4

I explore many different interpretations of the question of decision making. I begin with a survey of the views of prominent consequentialists. I then take a formulation of the question in terms of decision procedures ('Which decision procedure should I follow?') and analyse it in two stages. The first concerns the nature of decision procedures themselves. The second concerns the role in which they are to be assessed, or in other words, what it means to *follow* a given decision procedure. I conclude that the most natural analysis of the question involves the decision procedure that it is best for me to *be committed to*, rather than that which it is best for me to *comply with* or to *execute perfectly*.

Chapter 5

I apply the approach of Chapters 2 and 4 to the question of motivation, showing that the traditional forms of consequentialism cannot appropriately answer the question of motivation and then exploring the global consequentialist approach to the question. In particular, I examine the possibilities of focusing on individual motives, individual character traits, entire patterns of motivation, or entire characters, concluding that there are central moral questions in all of these areas, and thus that the question of motivation is less unified and more fragmentary than the question of decision making.

Chapter 6

I address several objections to the position that I have presented so far. Foremost among these is the objection made by Robert Adams, Brad Hooker and Bart Streumer that the decision procedure approach is inconsistent because it involves an obligation to follow a decision procedure even though it will lead us at some points to perform acts that we ought not do. I also address three other objections: that my approach involves a potentially damaging regress, that it would not remain distinctively consequentialist, and that it involves incorrect judgments about when we should hold attitudes of blame or guilt.

Chapter 7

I show how this consequentialist focus on decision procedures and motives can help to relieve the tensions between consequentialism, deontology and virtue ethics. In particular, I look at how the apparent disagreement between deontology, virtue ethics and act-consequentialism derives partly from their focus on different questions. Global consequentialism addresses all of these questions simultaneously and, I think, delivers intuitively acceptable answers. In this way, I show that the intuitions that lead some to deontology and virtue ethics can also lead to

a mature form of consequentialism.

Chapter summary

In this chapter I have set the stage for the rest of the essay. I first introduced act-consequentialism, explaining its criterion of rightness for acts, taking care to distinguish the objective, belief-relative, and evidence-relative versions of this criterion. I then presented an apparent objection: showing how motives of impartial benevolence and a strategy of naïve calculation could in fact be counter-productive. In response, I sketched the classical consequentialist reply: that we should instead have whichever pattern of motivation it is that leads to the best outcome, and decide in whichever way it is that leads to the best outcome. Finally, I outlined the structure of this essay, showing how I will develop this reply and the benefits that it can bring to the consequentialist program.

Chapter 2

Do the traditional forms of consequentialism suffice?

Over the next three chapters I shall explore how consequentialists can best formulate and answer the question of decision making, which is roughly phrased as 'How should I decide what to do?'. I shall focus on the approach suggested by many consequentialists: one should decide what to do in whichever way it is that leads to the best outcome. I have defined a decision procedure as a way of deciding what to do, so we can paraphrase this answer as saying that one should decide what to do by following the decision procedure that leads to the best outcome, or that the right decision procedure for someone to follow is the one that leads to the best outcome. Defenders of this answer thus have terms of moral approbation ('should', 'right') applied to the following of decision procedures. However, it is unclear how consequentialists can produce such a normative assessment of decision procedures. For example, if following a decision procedure is not itself an act, then the act-consequentialist criterion of rightness cannot be directly used to assess it.

In this chapter I shall first explore whether act-consequentialism has the resources to assess decision procedures, concluding that it cannot directly assess decision procedures in the appropriate way, but that it has some limited success providing indirect assessment. I shall then introduce rule-consequentialism and show that, while it can perhaps assess both acts and decision procedures, it does so at too great a cost to what I have been calling the fundamental consequentialist intuition. The ground will thus be prepared for considering newer consequentialist theories in the following chapter.

2.1 Act-consequentialism and decision procedures

In this section, I shall consider whether act-consequentialism can answer the question of decision making. Doing so requires a definition of act-consequentialism and throughout this essay I shall take it to be defined by the act-consequentialist criterion of rightness:

> An act is right *iff* it will lead to at least as much good as any alternative act

combined with an axiology to specify the goodness of outcomes.

Some people take act-consequentialism to be more than this. They take it to have principles (implicit or explicit) which can assess things other than acts, such as motives or beliefs. Indeed

I think that Bentham, Mill and Sidgwick were all committed to such additional principles, though they never made them explicit.[39] However, it has become traditional to take act-consequentialism to be defined by this criterion in conjunction with an axiology and it is useful to show that this minimal theory cannot satisfactorily address the question of decision making. In Chapter 3, I shall examine theories that are natural extensions of act-consequentialism and show how they can meet this challenge.

2.1.1 Decisions as acts

What, if anything, does the act-consequentialist criterion of rightness tell us about which decision procedure a person should follow or which decision procedure is the right one to follow? On the surface, it appears to talk only of the rightness of acts and to say nothing at all about the rightness of decision procedures. However, there are several reasons to believe that this might be mistaken.

Firstly, some authors have a very broad account of acts. For example, in *The Logic of Decision*, Richard Jeffrey says:

> '…where acts are characterized with sufficient accuracy by declarative sentences, we can conveniently identify the acts with the propositions that the sentences express. An act is then a proposition which is within the agent's power to make true if he pleases.'[40]

It is not quite clear from this whether Jeffrey means that all propositions which the agent can make true are acts or whether only a subset of these are. However, one of his examples is *we have red wine for dinner* which suggests that he means the former. An interpretation of the act-consequentialist criterion of rightness which used acts in such a broad manner would be able to assess decision procedures by looking at the 'act': *the agent follows such and such a decision procedure*. It would also be able to assess motives by looking at the 'act': *the agent is motivated in such and such a way*. I have no quarrel with such a theory of consequentialism (and indeed endorse something similar in Chapter 3); however it does go well beyond the ordinary understanding of the term 'act' and thus I shall not call it a species of act-consequentialism. In this essay I shall stick to the everyday interpretation of an act as something which someone *does* rather than as anything that an agent can make true.

An alternate approach to assessing decision procedures within act-consequentialism is through the consideration of mental acts. Suppose that Henry is sitting down and decides to stand up. In this case, he has done something even before he moves a muscle. He has made a decision, and this would appear to be a kind of act. It is not a physical act like standing up, or going for a walk, or even breaking a promise. Instead, it is a mental act, like imagining a perfect circle or counting to ten in your head. If someone accepts this, then they can say that since decisions are acts, act-consequentialism assesses them. Henry's decision to stand up would thus be right if and only if it (the *decision* to stand up) led to a better outcome than any

---

[39] See the quotations in Chapters 1 and 3 for more evidence of this.

[40] Jeffrey (1990), p. 84.

other act he could have performed.

Those who hold this view might go further and try to use act-consequentialism to assess the way in which the decision was made. There are several ways in which Henry could have made his decision: he might have accepted his first impulse, he might have determined that standing up was essential to getting himself a drink, or he might have calculated the benefits of standing up over remaining seated and found that the former won out. When we are talking of physical acts, we may consider different ways of doing something to be different acts. For example, if Michael can hit the ball hard or hit it softly, then we could say that he has (at least) two different acts available to him. Similarly, we might say that the different ways in which Henry could decide to stand up constitute different mental acts available to him. Thus we could say that Henry was right to decide to stand up by accepting his first impulse if and only if this was the best available alternative. If we accept all of this, we could say that act-consequentialism thus tells us the right way to decide what to do.

However, there are problems with this approach. There are difficulties involved in an account of mental acts and difficulties involved in individuating different ways of deciding to do a certain thing as different acts. There are also difficulties in having the act-consequentialist criterion comparing physical acts with decisions to do these acts and there are potential regresses concerning deciding to decide in a certain way.[41] These difficulties may be surmountable and it may in fact be possible to show that act-consequentialism is committed to assessing different ways of deciding what to do. However, there remains a greater problem.

The proposed act-consequentialist assessment of decision making is restricted to the consequences of a single decision. Many factors would go into determining the best way for Henry to make all of his decisions — the best single decision procedure to follow throughout his life — or his decisions over a period of his life, or in one domain of his life. However, on the interpretation above, act-consequentialism just determines the best way for Henry to decide this one thing — whether or not to stand up. Let us assume that the best outcome will be produced by his standing up, and that he has evidence for this. What is there to recommend one way of making this particular decision over another? Consider the purpose-built decision procedure: *if seated, stand up immediately*. If Henry decides to stand up via this decision procedure, he will stand up and will have wasted very little time in the decision making process. Indeed, if speed is of any importance in the situation, it seems that deciding to stand up using this decision procedure is superior to deciding to stand up according to a more generally useful (and thus more complex) decision procedure.

It might be thought that this aberrant decision procedure will be rejected when we look at the subjective forms of the criterion of rightness. However this need not be so. For example, the right act in the evidence-relative sense is the act that leads to at least as much expected good as any alternative, where the probabilities are given by the agent's evidence. If Henry's evidence suggests that outcomes where he stands up quickly are better than others, then deciding via the purpose built procedure above may well lead to more expected good than

[41] See, for instance, Smith (1991).

deciding via a slower, more general procedure and it will thus be the right way for him to decide what to do in the evidence-relative sense. While the evidence- and belief-relative senses of rightness refer to the expected good in their definitions, they do not require the agent to calculate this value, or even to consider it.

We are thus left with the problem that the act-consequentialist criterion of rightness as applied to ways of making decisions will often recommend the use of *ad hoc* decision procedures for the purpose of quickly producing the optimific physical action. Such a purpose-built decision procedure will beat all others in terms of its speed, simplicity and ease of use. However, this appears to have little relevance to the study of morality. What we were originally interested in was the best way of making decisions throughout one's life, or through some other significant period of time. This requires a generality and flexibility that is not captured when we consider only one particular decision.

One last possibility for this approach would be to consider act-consequentialism as applied to sequences of acts. Suppose the right act for Alice to perform is to cycle to the shop. This act of cycling to the shop is comprised of many smaller acts such as walking to her bicycle, unlocking the bicycle, climbing on the bicycle, and so on. If acts can be comprised of a sequence of smaller acts, then there is a sense in which act-consequentialism can assess sequences of acts as well as individual acts. Perhaps then, it can assess sequences of decisions?

Unfortunately, this does not seem possible. Even if act-consequentialism does assess certain sequences of acts in the above manner, it can only be applied to those which comprise some sort of unified whole. Unlike the sequence of acts involved in cycling to the shop, the sequence of acts that consists of all Alice's decisions throughout a part of her life has many temporal and causal gaps, and cannot plausibly be conceived of as a single act. It is thus not the type of thing which could be assessed with the act-consequentialist criterion of rightness, and the limitation to assessment of individual decisions remains.

2.1.2 Implicit obligations

An alternate account of how act-consequentialism can judge decision procedures comes from Gerald Lang.[42] He begins with a restatement of the act-consequentialist criterion:

(1) An agent ought to act in such a way as to produce at least as much good as that which would be produced by any other act available to her.

He then introduces a practical premise:

(2) An agent would not be able to produce acts which produce at least as much good as any other act available to her unless she is guided by some or other decision-making procedure, D^*.

[42] Lang (2004). His argument is originally aimed at act-utilitarianism, not act-consequentialism, and so I have simply substituted occurrences of the term 'utility' with 'good'.

And a deontic principle:

> (3) If an agent ought to secure *X*, and if doing *Y* is the only means of securing *X*, then the agent ought to do *Y*.

From these he draws the conclusion:

> (4) An agent ought to be guided by *D**.

I should note that this argument does technically go beyond act-consequentialism as I have defined it above, for it introduces an additional principle of instrumental rationality. However, since Lang is not considering this as an additional ethical principle, but merely as an essential part of practical rationality, I shall treat his argument as being about act-consequentialism and address it here.

The first thing to note regarding this argument is that Lang cannot strictly use (3) with (2) to produce his conclusion. In (3), Lang says 'if *doing Y* is the only means of securing *X*, then the agent ought to *do Y*'. The part of premise (2) that is supposed to take the place of 'do *Y*' is 'be guided by some or other decision-making procedure *D**', but is being guided by a decision procedure something that one can *do*? As discussed above, we may have reason to believe that making a particular decision in a particular way is an act, but in the sense that both Lang and I are interested in, being guided by a decision procedure is something that is extended over a great many acts. It does not seem plausible that 'being guided during many different parts of one's life by *D**' is something one can *do*.

This defect can be corrected if we change (3) to:

> (3a) If an agent ought to secure *X*, and if *Y* is the only means of securing *X*, then the agent ought to *Y*.

Here we have removed the word 'do' and thus made '*Y*' stand in for an entire verb phrase rather than a mere action. We could then properly substitute 'being guided by *D**' for *Y* to get:

> If an agent ought to secure *X*, and if being guided by *D** is the only means of securing *X*, then the agent ought to be guided by *D**.

This resolves the present difficulty, but it is worth noting that something substantive has been added to (3) in order to turn it into (3a). Whereas (3) could only be used to produce 'ought to do' claims, (3a) can generate claims like 'ought to be', 'ought to know' or 'ought to have'. These are essential for Lang's purpose as he wants to conclude that 'the agent ought to *be guided by D**'. However they take us qualitatively beyond the standard claims of act-consequentialism, which all apply to the rightness of acts, and the obligation to *do* acts. Lang is attempting to show that the combination of act-consequentialism with instrumental reasoning generates new obligations such as that of being guided by a certain decision procedure. This is an interesting proposal, but as we shall see, we will not need to come to a final assessment of whether it works, and whether (3a) is an admissible principle of practical

reasoning, for the argument suffers from more direct problems.

We now come to a dilemma for Lang's argument: either the conclusion cannot be reached due to the falsity of (2), or the conclusion must be interpreted in a manner that is much weaker than it first appears. Premise (2), as stated, is somewhat ambiguous. Does it refer to a single decision procedure being used to make all the acts or does it allow for different decision procedures for each choice of act? Let us first assume the former and restate the premise for clarity:

(2a) An agent would not be able to produce a sequence of acts, each of which is optimific, unless she is guided in all of them by some single decision-making procedure, D^*.

From here Lang might hope to get to the (more precise) conclusion:

(4a) An agent ought to be guided in her sequence of acts by some single decision-making procedure, D^*.

However, this cannot be achieved, since (2a) is false. Lang states that (2) 'trades on little more than the innocuous claim that, in order to act, we need to employ some or other decision-making procedure'[43], and while this claim may well be innocuous, (2a) requires much more. It requires that to act rightly on a whole sequence of acts, a single decision procedure must be used, but it is not clear why we should think so. Perhaps the first few right acts are chosen using one decision procedure, then the agent is convinced by a friend that an alternate decision procedure is superior and goes on to use that new decision procedure to perform the remaining acts. If this is possible (and surely it is), then (2a) is false. Lang's claim would only be innocuous if we were instead to interpret (2) as:

(2b) In any particular situation, an agent would not be able to produce the optimific act unless she is guided by some decision-making procedure, D^*.

This may well be true, and it can be used in conjunction with (3a) to reach a conclusion:

(4b) In any particular situation, an agent ought to be guided by some decision-making procedure, D^*.

However, in (2b) and (4b), the decision procedure does not need to be the same across a long sequence of choices that the agent makes. In a given situation, there are many decision procedures that would guide the agent and we could grant that she indeed ought be guided in her decision by one of them. However, at the next decision, there is another list of decision procedures that would appropriately guide the agent and she need not use the same one again. Thus, the conclusion (4b) would merely be saying that each time an agent makes a choice, she ought to be guided by one of the decision procedures that will lead to the right choice in that case.

[43] Lang (2004), p. 224.

We thus have another argument which suggests that act-consequentialism is committed to judging decision procedures to be right in particular situations (not a very useful thing in itself), but which does not judge their merits over a series of decisions.

2.1.3 Evaluation of decision procedures

Even if act-consequentialism cannot directly make claims like 'D^* is the right decision procedure for Michael to follow' or 'Michael should follow D^*', it may still be able to assess decision procedures in important ways. The next two subsections address this possibility.

Even if the act-consequentialist criterion of rightness cannot be applied to decision procedures, act-consequentialism can still assess decision procedures indirectly via its axiology. One could say 'the outcome of Michael's following decision procedure X is better than the outcome of his following decision procedure Y', or 'the outcome of Michael's following decision procedure Z is better than the outcome of his following any other decision procedure'.

This provides a kind of indirect evaluation of decision procedures. However, it doesn't directly answer the question of decision making ('What decision procedure should Michael follow?'), as it is logically consistent to claim that Michael should follow decision procedure Y even though his following decision procedure X would lead to a better outcome. This is closely analogous to the way that deontologists and rule-consequentialists can consistently claim that Michael should do act Y even though his doing act X would lead to a better outcome. Act-consequentialists must be consequentialists about acts, but nothing in the letter of the view stops them being nonconsequentialists about decision procedures. Thus, while the axiology does provide some indirect assessment of decision procedures, it does not go far enough.

2.1.4 The act of adopting a decision procedure

One reason for the historical focus on acts within consequentialism is that whenever we deliberately change the state of the world, we do so through acts. Thus, while our knowledge, beliefs, motives, decision procedures and attitudes may all be morally significant, if these are only ever deliberately changed via acts, then in some sense act-consequentialism can provide indirect assessment of all these features through its direct assessment of acts.[44]

For example, suppose that it would be good if Louise knew her times tables. We might informally express this by saying 'Louise should know her times tables'. However, suppose that the only way for Louise to come to know her times tables is by studying them and that this would guarantee she came to know them. If so, then studying the times tables would be a necessary and sufficient condition for knowing them and we could use the statement, 'Louise should study her times tables', to make both a direct normative claim about studying the

[44] This approach was originally suggested to me by Gustaf Arrhenius and Torbjörn Tännsjö at a seminar I gave in Stockholm. A similar approach can be found in Gruzalski (1986), p. 776.

times tables and a kind of indirect normative claim about knowing the times tables.

Such a method could also be applied to indirectly assess decision procedures. If it would be best for Michael to follow decision procedure *D*, and if the best way for Michael to follow *D* is for Michael to consciously adopt *D*, then we could say that Michael should adopt *D*. While we couldn't strictly answer the question of decision making by saying 'Michael should follow *D*', we could make the related claim about an act: 'Michael should adopt *D*'. We could then address the complaints about naïve calculation by making the following claims about acts: 'Michael should not adopt naïve calculation', or even 'Michael should adopt a decision procedure other than naïve calculation'.

This is an indirect way of dealing with the question of decision making, but it gives us much of what we sought. It focuses on assessing the deliberate transitions between outcomes, rather than assessing other features of the outcomes. It embodies the idea that once you know all statements of the form '*x* should do *y*', then in some sense you know everything you need to. There may be other non-deliberate transitions between outcomes, such as a tree growing or a person dying, but only the deliberate transitions — the *acts* — are typically thought to be subject to moral approbation.

However, there are times when the indirectness of the assessment causes problems. The first, and most obvious problem is if there is no act that is necessary and sufficient for making the agent begin to follow a particular decision procedure. In such cases — which are undoubtedly common — the connection between the act and the decision procedure becomes even less direct. There is also a serious problem concerning the time-lag between beginning to adopt a decision procedure and becoming fully committed to it. Consider the following example:

> *A change of pace*
>
> Andrew works as a doctor and currently follows decision procedure *X*, which is well suited to the particular demands of being a doctor and leads to better consequences than if he were to follow any other decision procedure. In one year's time he will retire from medicine and in his retirement it will be best for him to follow the more general decision procedure *Y*. However the process of changing his decision procedure is slow and difficult, so that if he attempts to adopt *Y*, this will take him a year of effort and in the meantime he will continue to be guided by *X*.

It is intuitive to say that over the coming year, Andrew should follow *X*. However, on the above approach we can only say that Andrew should adopt *Y*. The time-lag involved in adopting a decision procedure has let the claims about what is best to adopt and what is best to follow come apart.

It is my opinion that despite its drawbacks, this method offers the best way of normatively assessing decision procedures within act-consequentialism as defined here. It allows act-consequentialism to make claims about the adoption of decision procedures which are fairly closely connected to the intuitive claims of the previous chapter. Indeed I would reluctantly accept this approach were it not for the existence of extensions to act-consequentialism that

will be covered in the next chapter. I shall argue there that they provide much more precise renderings of our intuitive claims, greater theoretical simplicity and more illumination of the consequentialist project. These are strong claims and their evidence will have to wait for the next chapter. For now, let us put act-consequentialism to one side as a partial success in assessing decision procedures and see whether rule-consequentialism can do better.

2.2 Rule-consequentialism

We have seen that on its own, act-consequentialism does not have sufficient resources to directly assess decision procedures. It might be thought that rule-consequentialism offers a solution. In this section, I shall show that some forms of rule-consequentialism might be able to assess decision procedures, but only at the price of giving up the fundamental consequentialist intuition.

There are many ways in which rule-consequentialism can be formulated. A simple definition is given by Smart:

> '[Rule-consequentialism] is the view that the rightness or wrongness of an action is to be judged by the goodness or the badness of the consequences of a rule that everyone should perform the action in like circumstances.'[45]

It is said that act-consequentialism is a type of *direct consequentialism* because the rightness of an act is determined directly by comparing its consequences to the consequences of the alternative acts.[46] In contrast, Smart's formulation of rule-consequentialism is a type of *indirect consequentialism* because the rightness of an act is determined not by the consequences of the act itself, but by the consequences of a rule that everyone should perform the act in like circumstances. By allowing this layer of indirection, rule-consequentialism allows for a theory with consequentialist underpinnings that appears to generate more intuitive moral claims while avoiding the concerns regarding naïve calculation. However, it is well known that rule-consequentialism also introduces several difficult problems of its own.

A major criticism is that rule-consequentialism must either collapse into act-consequentialism or be incoherent. The first part of the argument can take several forms, but the basic idea is that the best rule will be to always do whichever act it is that leads to the most good. Opponents of rule-consequentialism then argue that if it somehow avoids such a collapse, it will be at odds with the fundamental consequentialist intuition of promoting the best outcomes, since it sometimes tells us that we should perform acts which lead to less good than the alternatives. More sophisticated formulations of rule-consequentialism, such as Brad Hooker's have therefore been introduced to attempt to address these (and other) criticisms:

---

[45] Smart (1973), p. 9.

[46] It is sometimes said (see Pettit and Smith (2000), and Kagan (2000)) that act-consequentialism is an indirect theory, assessing rules and other things in terms of the acts that they produce. I have instead been assuming that act-consequentialism assesses only acts and does so via its criterion of rightness. Nothing shall turn on this as it is merely a terminological issue.

> 'An act is wrong if and only if it is prohibited by a code of rules the acceptance of which by the overwhelming majority of people in each new generation would have the greatest expected value.'[47]

Hooker provides good arguments for why his version of rule-consequentialism does not collapse into act-consequentialism. Rule-consequentialism, on his account, is a theory which refers to rules that can form part of someone's actual decision processes. If we were to internalise the rule *always do what leads to the most good*, we would not always do what leads to the most good, but would presumably follow some form of naïve calculation. Given the well-known problems with this decision procedure, he rightly concludes that there would be other sets of rules whose internalisation would be less costly (by fitting more closely with common-sense morality) and which would lead to better a better outcome.

Rule-consequentialism, on Hooker's account, assesses two types of things: sets of rules, and acts.[48] The set of rules to be preferred is that set whose internalisation by most people in each new generation would lead to at least as good an outcome as the internalisation of any other set. Acts are then to be considered right if and only if they accord with this set of rules. While Hooker does not explicitly say that the preferred set of rules is the *right* set of rules, he does at times call it the *best* set of rules, or the 'ideal code'.[49] It is also a set of rules with a particularly close relationship to obligation, for we *ought* to act in accordance with this set of rules and (to rephrase this) it is the set of rules that we ought to comply with. Given this close moral approval for this set of rules, I shall follow Pettit and Smith in saying that rule-consequentialism does in fact offer a criterion of rightness for both rules and acts, where the right set of rules is that set whose internalisation leads to the best outcome and the right acts are those that comply with the right rules.[50]

Proponents of rule-consequentialism might thus say that it offers an account of how we are to decide what to do; one that is based on consequentialist principles and yet does not simply advocate the strategy of naïve calculation. Moreover, it explains this with two compatible forms of rightness: we should decide what to do by attempting to follow the right set of rules and (assuming we succeed in doing so) the acts we perform will be right.[51] We thus have an account which plausibly fulfils the aims of those act-consequentialists who wish to answer the question of decision making. However, I do not think that it ultimately succeeds.

Firstly, there is the question of whether rule-consequentialism really tells us to make our decisions by attempting to follow the right set of rules. It tells us which acts to perform, and

---

[47] Hooker (2004), §6. There is also a somewhat more complicated account in Hooker (2000), which refers to a specific axiology and includes rules for tie breaking.

[48] On other forms of rule-consequentialism it is individual rules, not sets of rules, which are the subject of moral assessment.

[49] See, for example, Hooker (2000), pp. 1–3.

[50] Pettit and Smith (2000). See also Kagan (2000). Nothing much will turn on this as by the end of the chapter I will reject rule-consequentialism despite this (charitable) assumption.

[51] This notion of 'attempting to follow' a rule is imprecise, but will do for now. We will investigate this more closely in Chapter 4.

which rules to comply with, but it doesn't necessarily tell us how to go about determining which acts to perform. If I chose my acts by rolling dice and happened to always comply with the rules, would I be open to censure from rule-consequentialism? To take this into account, rule-consequentialists may thus need to add a clause to their theory requiring us to make our decisions by attempting to comply with the optimific set of rules.

Secondly, there is a question of whether a set of rules is a sufficiently general approach to decision making. There are many ways in which we could decide what to do: many formulae we could invoke, many ways we could calculate and many ways we could rely upon certain intuitions. If deciding according to a set of rules could not include all such ways of deciding, then we would have a problem, for it may be that deciding according to one of these ways that is not captured by rules will lead to a better outcome than deciding in any rule prescribed manner. We would thus have a good reason to abandon the rule based approach in favour of a more general decision-procedure approach.

What are we to count as a rule? Hooker provides no precise definition, but appears to take a very general formulation. For example, he counts acting in accordance with naïve calculation as acting in accordance with a rule. Presumably the rule could be formulated as 'Always calculate expected utilities to such-and-such a degree of precision and then act upon your calculation'. If a rule is allowed to refer to the performing of calculations then it seems that it would be sufficiently general to accommodate all decision procedures, and I shall assume this to be the case for the remainder of this section. Even if it is somehow shown that there are decision procedures which cannot be accommodated, this is no major obstacle as one could always modify rule-consequentialism into some form of *decision-procedure-consequentialism*, where all references to the internalisation of sets of rules are replaced with internalisation of decision procedures.

Finally, a large difference between the standard consequentialist reply and rule-consequentialism lies in the way that sets of rules are to be judged. For rule-consequentialists a set of rules is judged on the basis of how good it is for it to be internalised by everyone or by the overwhelming majority of the population. Rule-consequentialists are motivated towards this position by its similarity to commonsense morality (and Kantian ethics) in prescribing a universal set of rules which we must all attempt to follow. There are also strong similarities to a legal framework in which one set of laws governs all.

However, we may well question why rule-consequentialism judges rules on the assumption that almost everyone will use them. Instead it seems that the obvious consequentialist answer to 'How should I decide what to do?' is that I should decide what to do in whatever way would be best, *given that everyone else does whatever it is that they actually will do*. We know that our decisions to change the way in which we decide things will not mysteriously compel others to decide in the same way, so why should we judge the set of rules (or decision procedure) as if they would? Furthermore, people differ greatly in their cognitive abilities, their willpower, and even their professions. Forcing them to use the same set of rules eliminates many potential benefits that could be gained by even roughly tailoring the decision procedure to

the person.[52] Rule-consequentialists have their reasons for their position and I am not attempting to argue conclusively against their view in this space. However, act-consequentialists who are merely seeking a way of assessing decision procedures are very unlikely to be swayed from their approach of assessing the consequences of an individual's choices in terms of what will actually happen.

Let us therefore consider a modified form of rule-consequentialism in which sets of rules are assessed with respect to individuals. Thus, we relativise the criterion of rightness for a set of rules so that it assesses whether a set of rules is right *for a given agent*, based on the consequences of that given agent's internalising those rules. An act by an agent is then said to be right whenever it is in compliance with the rules that are right for that agent. This is a rather major modification of the theory (and one that would not typically be supported by the advocates of rule-consequentialism), but it successfully defuses the previous objection.

Does this modified theory provide us with a good way of assessing acts and decision procedures? I believe that it still does not. On this modified view, rule-consequentialism is committed to the claim that it is right for an agent to perform a given act if and only if it is not forbidden by the set of rules which it is right for her to internalise. However, this means that the right act will often not be the one that leads to the best outcome.

Suppose, for instance, that the best set of rules for Jim to follow includes a rule stating that he should always tell the truth. There might be various exceptions to this built in to cover disastrous consequences, but we can assume that to lie in a fairly typical situation (when it would only lead to a small increase in the good) will be forbidden. Consider such a situation in which the best outcome can only be achieved by Jim lying, where the benefits of this are rather small, and where Jim is aware of these facts. If we share the fundamental consequentialist intuition that what matters is the goodness of the outcome, then it seems that — contrary to any of these forms of rule-consequentialism — it is right for Jim to lie.

While this modified form of rule-consequentialism allows consequentialists to assess the way we make decisions, it does so at the cost of violating one of the major motivations for consequentialism itself. Hooker is correct in holding that this does not show the various versions of rule-consequentialism to be incoherent, for one might find it more important that a moral theory fits our pre-existing intuitions than that it respects the fundamental consequentialist intuition. However, this failure constitutes a sufficiently large flaw to make act-consequentialists who are looking for a criterion of rightness for decision procedures look elsewhere — especially given the existence of a much more favourable alternative.

---

[52] There would, of course, be benefits to everyone sharing a common decision procedure in terms of predictability and perceived fairness; however if this is a true motivation for the rule-consequentialists, then the question of whether we should use a common decision procedure should be left as an empirical one: 'Do the benefits actually outweigh the costs?'

Chapter summary

In this chapter I showed that neither act-consequentialism nor rule-consequentialism can satisfactorily formulate or answer the question of decision making. I first considered act-consequentialism, showing that treating acts as decisions does not help and nor does Lang's proposal to supplement act-consequentialism with an additional deontic principle. I then considered whether act-consequentialism could indirectly assess decision procedures and showed that it had some very limited success by using its axiology to assess outcomes in which a decision procedure is followed, and some greater, though still limited, success by assessing the act of adopting a decision procedure. Finally, I considered rule-consequentialism and showed that even if it is construed in such as way as to allow it to individualistically assess the following of decision procedures, this still comes at the cost of violating the fundamental consequentialist intuition when it comes to the assessment of acts. This sets the stage for theories that assess all focal points (including acts and decision procedures) in terms of their consequences.

Chapter 3

Applying consequentialism to everything

The question of decision making is 'How should I decide what to do?'. We have seen that act-consequentialism cannot directly answer this because it does not assess decision procedures. Some versions of rule-consequentialism assess decision procedures, but they provide only indirect assessment of acts, leading to conflicts with what I call the fundamental consequentialist intuition: that outcomes be as good as possible. What we need is a theory that assesses both acts and decision procedures, and yet preserves the fundamental consequentialist intuition by assessing each directly in terms of the outcome it would lead to.

Over the last few decades, several theories have been proposed which do precisely this and the broad approach has come to be known as global consequentialism. In this chapter I shall present all six major versions of global consequentialism, breaking them down into their constituent components and exploring the space of possible global consequentialist theories. I shall also construct two new global consequentialist theories, which I believe represent the most promising lines of development. Either of these theories will suffice to make the assessments of acts, decision procedures and motives that are needed to answer the questions of decision making and motivation that I posed in Chapter 1.

This a long chapter, and at times it is quite technical. However, this is necessary as it forms the core of my project, providing the theoretical underpinnings that will be used in the later chapters to assess decision making and motivation. Because of the importance of this topic in both this essay and in the study of consequentialism at large, I have endeavoured to provide a thorough treatment of the existing theories and of the possibilities for new theories.

3.1 Evaluative focal points

Acts and rules are two examples of what Shelly Kagan calls *evaluative focal points*.[53] They are the things that receive the direct consequentialist assessment in act-consequentialism and rule-consequentialism respectively. However, they are not the only possible targets for assessment in consequentialist theories. For example, Robert Adams has advocated a form of motive-consequentialism in which an agent's motives are directly assessed in terms of the

[53] Kagan (2000).

good they would lead to.[54]

Adams is not alone in wanting to directly assess a person's motives in terms of their consequences. Indeed, Bentham, Mill and Sidgwick all made claims of this type. For example, Bentham's *Introduction to the principles and morals of legislation* states:

> 'With respect to goodness and badness, as it is with every thing else that is not itself either pain or pleasure, so is it with motives. If they are good or bad, it is only on account of their effects: good, on account of their tendency to produce pleasure, or avert pain: bad, on account of their tendency to produce pain, or avert pleasure.'[55]

In Chapter 1, we saw that Mill and Sidgwick both clearly advocate a consequentialist assessment of motivation:

> 'But I hold that the very question, what constitutes this elevation of character, is itself to be decided by a reference to happiness as the standard.'[56]

> '…if experience shows that the general happiness will be more satisfactorily attained if men frequently act from other motives than pure universal philanthropy, it is obvious that these other motives are reasonably to be preferred on Utilitarian principles.'[57]

Like Bentham, Sidgwick also explicitly suggested that this assessment of motives in terms of their consequences could be extended to other objects of evaluation:

> 'While yet if we ask for a final criterion of the comparative value of the different objects of men's enthusiastic pursuit, and of the limits within which each may legitimately engross the attention of mankind, we shall none the less conceive it to depend upon the degree in which they respectively conduce to Happiness.'[58]

In addition to acts, rules and motives, there are many other natural focal points for moral theories. For example, we may include norms, character traits, decision procedures, institutions, and even entire lives. For each of these focal points, one could construct a consequentialist theory in the mould of act-, rule-, or motive-consequentialism: directly assessing the one central focal point and either providing some form of indirect assessment, or else remaining silent regarding each of the others.

Alternatively, we could directly assess each of these focal points in terms of the good that they would lead to. This approach seems more in keeping with the above quotations by the

---

[54] Adams (1976). On the strict version of Adams' theory acts are not assessed at all, though he is open to the possibility of holding motive-consequentialism in conjunction with a theory that assesses acts. Other authors have suggested versions of motive-consequentialism in which the right acts are determined indirectly by what a person with the right motives would do.

[55] Bentham (1789), p 102.

[56] Mill (1843), Bk. VI, Ch. 12, §7.

[57] Sidgwick (1907), p. 413.

[58] Sidgwick (1907), p. 405.

classical utilitarians and has been advocated explicitly by a number of prominent consequentialist thinkers.[59] It may also be implicitly accepted by many act-consequentialists, as it is not in conflict with the act-consequentialist criterion of rightness. Instead, it can be seen as an extension of act-consequentialism, allowing the assessment of new focal points about which the theory was formerly silent. This approach to consequentialism is both new and old: new in the sense that it was first explicitly formulated in the 1980s, and old in the sense that it may have been implicitly accepted by even the classical utilitarians.[60]

The advocates of this approach do not limit themselves to the list of focal points above. Instead, they take the approach to its natural conclusion and assess *everything* directly in terms of its consequences. I shall thus follow Pettit and Smith in calling this approach *global consequentialism*, and using the term *local consequentialism* to refer to approaches like act-, rule- and motive-consequentialism which restrict their direct consequentialist assessment to a single focal point.[61] Even if we set aside questions about axiology (i.e. the goodness of individual outcomes), there are many forms of global consequentialism. I will therefore use the term to describe the general class of theories which apply a consequentialist criterion to everything, rather than to designate any specific theory within this class.

The potential advantages of global consequentialism are many. The addition of new focal points increases the expressivity of consequentialism. It allows consequentialists to directly answer questions about the best motives, the best system of government, or the best way to decide what to do. This also allows consequentialists to bring consequentialism into the strongholds of the deontologists and virtue ethicists: it provides an explanation of the importance of character and of rules of conduct, going so far as to show how many acclaimed virtues and rules have systematically led to good consequences, and even daring to suggest ways in which they might be improved.

It promises to do all of this with very few additional drawbacks. Unlike rule- or motive-consequentialism, it never departs from the core consequentialist ideal of assessing things by their consequences. By encompassing all focal points, rather than just the most prominent ones, it maximizes its expressive power while remaining non-arbitrary. Indeed, some systems of global consequentialism promise to be simpler than act-consequentialism itself, for by allowing everything to be morally assessed, they no longer need an associated theory of acts.

---

[59] In addition to those that I shall discuss in this chapter, it is explicitly advocated by David Brink: 'The teleologist can assess motives, rules, and institutions as well as actions. These too the teleologist will assess by the value they realize.' [Brink (1986), p. 421], and by Julia Driver: 'it is the view that the moral quality of a person's action (character trait, etc.) is determined by its consequences.' [Driver (2001), p. xiv].

[60] The historical focus on *act*-consequentialism as opposed to global consequentialism may have been a by-product of the formulation of rule-consequentialism. Rule-consequentialists used the term 'act-consequentialism' to refer to their rival theory and this term may have led to the idea that acts were to be the key focal point for direct consequentialists, rather than just one focal point among many.

[61] Pettit and Smith (2000).

3.2 Four rough accounts of global consequentialism

3.2.1 The four accounts

So far, there have been six major accounts of global consequentialism: those given by Parfit, Railton, Kagan, Pettit and Smith, Feldman, and Crisp.[62] The last two in this list have a distinctly different character from the first four, and they warrant separate treatment. I shall thus leave Feldman and Crisp until section 3.3 and begin by addressing the other four accounts.

Though Parfit, Railton, Kagan, Pettit and Smith all advocate forms of global consequentialism, they do so only in general terms: showing why we need to go beyond the various forms of local consequentialism, but without explaining exactly how this is to be done. While their main points are clear, there are nuances on which they do not elaborate. Since their explanations of global consequentialism are also very concise, it will be best to simply quote them each in full rather than trying to summarise.

In *Reasons and persons*, Parfit introduces his version of consequentialism:

> 'There are different versions of *Consequentialism*, or *C*. C's *central claim* is:
>
> (C1) There is one ultimate moral aim: that outcomes be as good as possible.
>
> C applies to everything. Applied to acts, C claims …
>
> (C2) What each of us ought to do is whatever would make the outcome best.
>
> … Consequentialism covers, not just acts and outcomes, but also desires, dispositions, beliefs, emotions, the colour of our eyes, the climate, and everything else. More exactly, C covers everything that could make outcomes better or worse. According to C, the best possible climate is the one that would make outcomes best. I shall use 'motives' to cover both desires and dispositions. C claims
>
> (C5) The best possible motives are those of which it is true that if we have them, the outcome will be best.
>
> As before, 'possible' means 'causally possible'. And there would be many different sets of motives that would be in this sense best: there would be no other possible set of motives of which it would be true that, if we had this set, the outcome would be better. I have described some of the ways in which we can change our motives. (C2) implies that we ought to try to cause ourselves to have, or to keep, any of the best possible sets of motives.'[63]

---

[62] Parfit (1984), pp. 24–28, Railton (1988), Kagan (2000), Pettit and Smith (2000), Feldman (1986, 1993), Crisp (1992).

[63] Parfit (1984), p. 24–26.

In his paper, 'How thinking about character and utilitarianism might lead to rethinking the character of utilitarianism' Peter Railton proposed a theory that he called *valoric utilitarianism*.[64] This theory differs from act-utilitarianism in two ways. The first is that it is not directly concerned with either rightness or goodness. Instead, Railton defines an act to be more *morally fortunate* than another if it leads to the promotion of more non-moral value. To this, one could add a theory of rightness and goodness which need not be a maximising theory. The idea is that the combined theory might be able to make consequentialist-style judgments concerning the need for maximization (using the language of moral fortunateness), and yet reserve the everyday terms rightness, wrongness, blameworthiness to bear meanings closer to their commonsense use.

However, it is the second departure from act-utilitarianism that is most relevant here:

> '…valoric utilitarianism has direct application not only to acts, but to any object of moral assessment. In this way it differs not only from familiar indirect utilitarianisms, but also from direct act utilitarianism. One can ask how morally fortunate an individual act is, or how morally fortunate actions of that kind usually are, or how morally fortunate it would be if everyone regularly took such actions, and so on. And one may ask how morally fortunate it is that on a given occasion an individual possesses or acts from a given character, or how morally fortunate a character of that kind usually is, or how morally fortunate it would be if most people had such characters, and so on.'[65]

In 'Evaluative focal points', Kagan argues forcefully against all forms of indirect consequentialism before introducing his solution:

> '…if there is a plausible version of consequentialism, it will evaluate *both* focal points — acts *and* rules — directly. Neither focal point will be elevated to the status of primary evaluative focal point; neither focal point will be evaluated only indirectly…. I have restricted our discussion to two focal points: acts and rules. But as I have already suggested, there are many other evaluative focal points that have been endorsed as *primary* focal points — such as motives, norms, institutions, and decision procedures. As I see it, the most plausible version of consequentialism will indeed be direct with regard to *all* of these. In fact, once we free ourselves from the thought that the evaluative focal points must be at least prima facie plausible candidates for the office of *primary* focal point, we realize that absolutely every kind of thing is a potential evaluative focal point (atoms, the weather, sewer systems, suns). So I believe that the most plausible version of consequentialism will be direct with regard to everything.'[66]

In the same volume, Pettit and Smith introduce their own theory of global consequentialism. Like Kagan they spend most of the article arguing against indirect consequentialism and only a few paragraphs explaining their alternative:

---

[64] Railton (1988).

[65] Railton (1988), p. 410.

[66] Kagan (2000), pp. 150–151.

> 'Global consequentialism identifies the right *x*, for any *x* in the category of evaluands — be the evaluands acts, motives, rules, or whatever — as the best *x*, where the best *x*, in turn, is that which maximises value.… So, for example, according to global consequentialism, the right act for someone to perform is the act that has greater value than any of the acts that might have been performed instead; the right motive-set for someone to have is the motive-set whose possession has greater value than any of the motive-sets that might have been possessed instead; the right set of rules for someone to have internalised is that set of rules which has greater value than any of the sets of rules that they might have internalised instead, and so on.'[67]

The above accounts have many similarities. In particular, they all advocate that 'everything' should be assessed in terms of its consequences. However, there are also a number of differences and ambiguities. By analysing these, we shall be better able to survey the space of possible global consequentialist theories.

3.2.2 Focal points and roles in which to evaluate them

To assess a particular act using act-consequentialism, we examine the outcome that would occur were it performed. Let us say that the *role* in which act-consequentialism assesses an act is its *being performed* — as opposed to its being imagined, or its being denounced. It is possible to set up act-consequentialism using different conceptions of acts, and these need to be assessed in slightly different roles. For example, we could consider acts in a narrow sense, such that each act can only be performed by one particular person, and only at one particular time (such as the act of *Henry standing up at 12pm*). In this case, we don't need to know any more than that the act is performed. Alternatively, if we consider acts in a broad sense, such that two people could be said to perform the very same act (such as the act of *standing up*), then it would no longer be enough to know that the act was performed. In order to know what outcome the act would have, we would need to know who performed it and when. In this case the role in which to assess the act would be that of *being performed by a given person at a given time* (or over a given period). In either case, the appropriate role is obvious and rarely mentioned explicitly. However, when we consider other focal points, the choice of role becomes more complicated.

When assessing sets of desires, dispositions, beliefs and emotions there seems to be a simple common role. We can look at the outcome that would result from a given agent possessing the given set of desires/dispositions/beliefs/emotions at a specified time (or over a given period of time). Here the term 'possessing' has a subtly different meaning for each of these focal points, but in each case the correct interpretation seems clear.

However, what are we to say about assessing a given institution or climate? These are not things that an individual can perform or have. In the case of climate, Parfit writes that 'the best possible climate is the one that would make outcomes best'[68], but in what role are we to

---

[67] Pettit and Smith (2000), p. 121.

[68] Parfit (1984), p. 25.

evaluate this climate? Are we to consider it as the prevalent climate for Oxford? for England? for the entire world? It is not clear. Moreover, whichever answer we choose, we exclude the other ways of assessing the climate and this seems unnecessarily limiting.

Perhaps we should make the region affected by the climate a parameter in our assessment. We could then assess a climate in terms of the outcome that would result from a given region having the climate over a given period of time. This would work, but it might make us wonder whether there is a single best role in which we are to assess each class of things. With the exception of Railton, all the above authors seem to assume that there is, for they provide examples of things being assessed in certain roles without specifying how we are to know which roles to use.

A different approach would be to make our assessments relative to a particular role. Thus we could assess a climate in the role of being had by Oxford in the 18th Century, or being had by Madagascar throughout its existence. This conception of roles is philosophically simple: a role is a single place predicate.[69] When given a token from the class under consideration, it forms a proposition. For example, the first role above is the predicate: *is the climate had by Oxford in the 18th Century*. When a particular climate, such as *temperate* (or something much more specific) is provided, this forms the proposition: *temperate is the climate had by Oxford in the 18th Century* or to rephrase this: *Oxford had a temperate climate in the 18th Century*. We can then assess this proposition in terms of the outcome that would occur were it to be true.

There are three major virtues of the above approach. We have seen two: it avoids the arbitrariness that would otherwise exist in specifying which role we should use to assess tokens from each class and it allows us to use the familiar philosophical machinery of predicates and propositions to make evaluations. The third virtue is that of increased expressiveness.

For example, we could already assess the best set of motives for Jane to have, but now we can also assess the best set of motives for Jane to inculcate within herself, the best set of motives for Rochester to inculcate in Jane, the best set of motives for Jane to inculcate in Rochester, the best set of motives for Jane to aspire to, and the best set of motives for Jane to renounce. Without treating the role as a parameter of the evaluation, we would not be able to express all of these.

This approach is in conflict with a claim made by Pettit and Smith:

> 'The intuitive notion of rightness, as applied to acts, is thus just the idea of the act that it is right for someone to perform – as opposed, say, to the act that it is right for someone to forget. In the latter case the evaluand is forgettings, not acts.'[70]

---

[69] The term 'predicate' can be used to refer to a linguistic expression such as 'is red' or to the relation denoted by this expression (in this case: *is red*). I use it in this latter sense.

[70] Pettit and Smith (2000), p. 133. Note that Pettit and Smith are not explicitly denying the 'role' approach that I favour, as it is not clear that they are aware of this possibility.

However, I don't think that Pettit and Smith's alternative can really capture all the same assessments. For example, the best 'forgetting' may be of a certain dream, or visual experience and not of an act at all. There would thus be no way to capture the idea of the best *act* to forget. The additional expressivity comes from considering all combinations of focal points with roles in which they could be considered.

While this claim by Pettit and Smith is in conflict with the role approach, is should not be taken as actively denying this method, for it is doubtful that they were explicitly aware that such an approach was possible. Parfit also seems to use one role per focal point, though he too does not explicitly consider the alternative. In contrast, Kagan explicitly mentioned the possibility of assessing focal points in different roles:

> 'But rules … don't generate results in and of themselves. Rather, they generate results only when they are 'embedded' in some way: they generate results when they are thought about, taught, accepted, disdained, mentioned, mocked, acted upon, flouted, or what have you. So talk of selecting the rules that would have the best results is necessarily shorthand for talk of selecting the rules that would have the best results *if* they were embedded in some specified way.'[71]

However, he wrote this in the context of a rule-consequentialist choosing just one of these ways in which to assess a set of rules. In his discussion of global consequentialism towards the end of the article, he did not re-emphasise this point and it is thus unclear whether he expects the global consequentialist to argue over which role (embedding) is appropriate for which focal point, or to assess focal points relative to any specified role.[72]

Only Railton explicitly advocated assessing a single focal point in multiple ways. However, he was not clear about what roles are available or how exactly we are to assess things within them. For his each of his two example focal points — acts and characters — he considered three ways in which they could be assessed:

(1) by the outcome of their being performed/possessed on the occasion at hand,

(2) by the 'usual' outcome of that type of act/character being performed/possessed,

(3) by the outcome of their being performed/possessed 'regularly' or 'by most people'.

(1) is a very standard assessment and can be covered on all global consequentialist theories. (3) is akin to the type of universalisation considered by rule-consequentialism or Kantian ethics and goes beyond the examples of roles discussed so far. However, there is nothing that stops the role-based approach from making such an assessment. For example, we could ask how good it would be for a specified character to be in the role: *is a character possessed by most people*, or for a specified act-type to be in the role: *is a regularly performed act-type*. Thus, a role-

[71] Kagan (2000), p. 137.

[72] One thing that points in favour of the latter is his choice of atoms as a focal point. What could be the canonical role for atoms?

based global consequentialism can simultaneously make judgments about various forms of universalisation alongside judgments on individual instances.

However, (2) is more challenging to accommodate within this framework. It does not look at the outcome of something being in a certain role, but at the 'usual' outcome of it being in that role. It is thus best thought of in a similar way to evidence-relative or belief-relative judgements. We can accommodate belief-relative judgments in role-based global consequentialism by asking for the expected goodness of a focal point being in a given role (where the probabilities are given by an agent's degrees of belief). If we have a particular definition of 'usual outcome' in mind, we could formulate a way of specifying the 'usual goodness' of something being in a given role and then ask what is the usual goodness of a specified character being possessed by a person. Later in this chapter we shall come back to the topic of actual goodness versus belief-relative or evidence-relative goodness, but since we shall have no particular need of 'usual goodness' I shall not develop it further.

In summary, we have two approaches for assessing a focal point. We can do so relative to a fixed focal-point dependent role and several parameters (such as an agent, a time, a period of time, a region etc), where the list of parameters may itself depend on the focal point. Alternatively, we can assess the focal point relative to an arbitrary role which incorporates the relevant agents, times, or regions in question. This latter option is more expressive and I prefer it, though there is room for disagreement.

3.2.3 Right or best?

A key detail in any global consequentialist theory is whether it assesses its focal points in normative terms (right, wrong, should, ought…), or in merely evaluative terms (good, bad, better, best…).[73] Pettit and Smith make all of their global consequentialist claims in terms of rightness: 'the right act', 'the right motive-set' and 'the right set of rules'.[74] In contrast, Parfit's assessment of non-act focal points are all couched in evaluative terms: 'the best possible climate' and 'the best possible motives'. He reserves the normative terms for acts: 'what each of us ought to do'.[75] Though it is less clear, Kagan seems to follow Parfit, mentioning 'the best rules' and 'the best motives', but 'the right thing to do'.[76] Finally, Railton abandons all claims of rightness *and* goodness in his global consequentialist theory.

Let us call those theories that apply normative terms to everything *normative global consequentialism* and those theories that restrict normative terms to acts *semi-normative global consequentialism*. In the context of global consequentialism, normative and evaluative terms are not on an equal footing. As consequentialists, all of these philosophers would presumably

[73] Some people reject this way of making the normative/evaluative distinction. I am not saying that the present distinction is the best use of these terms, but am merely exploring a potential source of conflict between global consequentialist theories.

[74] Pettit and Smith (2000), p. 121.

[75] Parfit (1984), pp. 24–26.

[76] Kagan (2000), p. 152.

accept that if something is right, it is also best. However, movement in the other direction (from the evaluative to the normative) is contentious for focal points other than acts: despite going out of their way to stress the equal status of non-act focal points, Kagan and Parfit still stop short of extending normative terms to cover them.

It is not clear how important this distinction is. Parfit has since explained that his restriction of normative terms to acts is not meant to be substantive.[77] He calls an act 'right' and a motive 'best' merely because he thinks that talk of the 'right motive' is less natural than the 'best motive'. His choice thus reflects perceived irregularities in the English language rather than perceived irregularities in the structure of ethics. It is not clear whether Kagan shares such a view, or whether he holds that refusing to apply normative terms to non-acts reflects an important philosophical distinction.

In any event, some philosophers will see this as a substantive point and so I provide an analysis of the issues involved. However, since the choice between normative and semi-normative global consequentialism turns out to not make much difference to the main points of this essay, I have placed this analysis in an Appendix. In summary, I find that there is an open question as to whether we can normatively assess things other than acts. If we can't, then semi-normative global consequentialism provides a satisfactory answer to the questions of decision making and motivation, and normative global consequentialism is impossible. If we can, then we should use a version of normative global consequentialism.

3.2.4 Formalizing role-based global consequentialism

We can now take the above considerations into account and formulate principles for a role-based theory of global consequentialism. We shall start with a principle of goodness:

Global consequentialist formula for objective goodness within a role

the objective goodness of x being in role R = the goodness of the outcome that would occur were x in role R

For example, the goodness of Jane being motivated in a certain way is the goodness of the outcome that would occur were Jane motivated in that way. The principle presented here is an objective one, but it is easy to produce similar accounts for the various subjective conceptions of goodness. For example:

Global consequentialist formula for belief-relative goodness within a role

the belief-relative goodness of x being in role R = the expected goodness of x being in role R, where the probabilities are given by the agent's degrees of belief

There is, however, a serious complication with both these definitions of goodness. They provide values of goodness which take the whole world into account. For example, if I

[77] Personal communication.

actually have a cup of coffee this morning, then the objective goodness of my doing so is the goodness of the actual world, and the belief-relative goodness is very close to this. When we wonder how good it is that I have a cup of coffee this morning, we are not typically interested in the goodness of the entire world where that occurs, but in the degree to which the outcome in which I have a cup of coffee is better than the outcome in which I don't.

Can we therefore devise a superior definition of goodness by comparing something being in a role with it not being in that role? Unfortunately, we can't. As Norcross points out in 'Good and Bad Actions', the way in which we judge the goodness of actions (and other focal points) depends upon the context.[78] In some cases we compare someone's performance of an action to whatever they would have done if they hadn't performed it, while in other cases we compare the performance of an action to complete inaction. The context determines which of these we use and there is no simple definition that suffices to capture this ambiguity in how we assess the goodness of actions. Norcross thus abandons talk of goodness of actions in favour of betterness. I shall mostly follow suit: warning readers against using the above definitions for intuitive assessment of goodness, but retaining the definitions in a purely technical capacity as they can be legitimately used in fleshing out the concept of betterness within a role.[79]

Global consequentialist criterion of betterness within a role

x is better than y in role R *iff* the goodness of x being in role R is greater than the goodness of y being in role R

I have presented this as a single criterion, but it is really a family of criteria. If we take the goodness on the right hand side to be objective goodness, then we get an account of objective betterness; if we take it to be belief-relative goodness, then we get an account of belief-relative betterness and so forth. This can be interpreted in two ways. If you are a pluralist about these senses of goodness/betterness/rightness then you can accept all members of this family as useful criteria. If you only accept one of them, then you can focus on that sense and ignore the others. For the rest of this chapter I shall continue to present goodness/betterness/rightness in this neutral way, avoiding the need to discuss them separately.

We can also add another principle to talk about the degree to which something is better than another thing in a given role. Thus, even though we will not speak intuitively of the goodness of something being in a role, we can still speak quantitatively and respect the fact that some differences in value are much greater than others:

[78] Norcross (1997).

[79] One way to make this distinction would be to use a deliberately technical term such as 'world-goodness' for the quantity specified above. Then this term could be used in the subsequent formulae and the theory would make no judgments of goodness *simpliciter*.

Global consequentialist formula for degrees of betterness within a role

x is better than y in role R to the degree to which the goodness of x being in role R is greater than the goodness of y being in role R

Unfortunately, there is a further complication which must be dealt with. Whereas acts are often defined in such a way that two alternative acts cannot both be performed, this is not true for focal points in general. Such a situation can lead the above principle to make unintuitive judgments. For example, suppose we are considering two events — *the holocaust* and *the scratching of my finger* — within the role of *happening*. Now suppose that I don't actually scratch my finger. In this case, the outcome were the holocaust to happen is just the actual historical outcome, while the outcome were I to scratch my finger is an outcome that is slightly worse than the actual outcome (for it includes both the holocaust and the scratching of my finger). Thus, if applied to this case, the above principle would rate the scratching of my finger as worse than the holocaust.

The problem stems from the two things not being proper alternatives for the role: one is in the designated role whether or not the other is. To get around this, we could make any of the following stipulations:

(1) The criterion can only be applied when it is impossible for x and y to both be in R

(2) The criterion can only be applied when, were x to be in role R, y would not be and were y to be in role R, x would not be.

(3) Roles must be specified such that no two things can simultaneously fill them.

Stipulation (1) restricts betterness judgments to mutually exclusive things, and thus avoids the problem. Stipulation (2) also avoids the problem, but does so without sacrificing as much expressiveness. For example, it lets us compare two events (both of which could happen) so long as were one to happen, the other wouldn't. Both (1) and (2) are ways of saying that x and y must be *alternatives*. We could say that (1) restricts x and y to being *hard alternatives*, while (2) restricts x and y to being *soft alternatives*. Stipulation (3) is quite similar to (1), but puts the burden on the specification of roles. For example, it wouldn't allow the role of *being a motive possessed by Susan*, as there are outcomes in which Susan possesses more than one motive, but it would allow *being the motive-set possessed by Susan*. Since (2) is the most expressive, it is my preferred option.[80]

It is also worth noting that the global consequentialist criterion of betterness involves a counterfactual and thus requires some explanation of the type of possibility involved. This can be more liberal than the type of possibility involved in 'ought' implies 'can', since we may want to compare two things even if no-one can bring them about. This could even go beyond physical possibility. For example, one could ask 'Would it be better for Alex to have

[80] This problem was also noticed by Norcross (*ibid.*) in the context of comparisons between acts made at different points in time. His conclusion was to abandon all comparison between acts that are not from the same alternative set, which is a version of (1).

the ability to fly or the ability to become invisible?' Even though both abilities might be physically impossible, we can still make sense of this question relative to a shared background of how such supernatural abilities are supposed to work. By using such a liberal kind of possibility in the betterness judgments, we maximise the expressiveness of the theory.

We can now add a principle to say when something is the best within a role:

Global consequentialist criterion of best within a role

x is the best f in role R *iff* for each other f, x is better than it in role R

The 'f' in the above principle stands for any particular focal point (such as *acts* or *institutions*), whereas the 'x' is a member of this category (such as *sitting down* or *the House of Lords*). For example, one might use the criterion to ask whether the House of Lords is the best institution for legislative review in England. While we defined betterness without reference to the focal point, it is needed here to provide a class of alternatives with which to compare x.

Note also that since this principle uses betterness comparisons between x and every other instance of the focal point, it requires that the instances of the focal point are all alternatives in one of the above senses. Thus, the criterion tells us whether a certain pattern of motivation is the best one for someone to possess since patterns of motivation are alternatives to each other, but not whether a certain event is the best event to happen, because there are possible events which are not alternatives. In general, this reluctance to make judgments on such strange questions seems to be a virtue of the theory rather than a flaw.

This principle is very similar to the formulations of Parfit and Kagan, although it explicitly allows focal points to be assessed with respect to different roles. Together, the two principles we have formulated provide a basis for making evaluations, but do not allow us to make any normative claims (not even for acts). We thus have a theory of merely evaluative global consequentialism, or 'scalar global consequentialism'. If we want a theory of semi-normative global consequentialism, we can add the act-consequentialist criterion of the rightness of acts (from Chapter 1). This would match Parfit and Kagan's approaches, and would provide a specific theory of role-based semi-normative global consequentialism.

To form a theory of normative global consequentialism, we could add the following principle instead of the act-consequentialist criterion of rightness of acts:

Global consequentialist criterion of rightness within a role

x is the right f in role R *iff* x can be in role R;
for each other f that can be in role R, x is better than it in R;
and it isn't better to have nothing in role R

Adding this principle gives us a version of normative global consequentialism that is similar to Pettit and Smith's, but also allows for assessment with respect to arbitrary roles.

It is important to note that in this principle, the conception of possibility given by the word 'can' must be much more limited than those discussed earlier. While we may sensibly talk about 'the best magical power for someone to have' even though any magical power would

be a physical impossibility, it would be another thing altogether to talk about 'the right magical power for someone to have', for this would presuppose that they *can* have a magical power in a sense relevant to the injunction that 'ought' implies 'can'.

We might therefore imagine that we are dealing with an individualistic and physical sense of can, such as '*s* can do *x* if and only if were *s* to desire to do *x*, *s* would do *x*'. However, there are sensible sounding rightness claims that go slightly beyond this level of possibility. For example, that 'the current climate is the right climate for the world to have in 2100'. It may be that there is no change to an individual's desires which would reverse climate change, but a change to multiple individuals' desires would suffice.

In its current form, this principle is somewhat ambiguous between the individual conception of possibility and the collective conception: if we used it to produce a claim about the right act for someone to perform then we would naturally assume a conception of possibility relative to that person, while for claims about the future climate we would naturally assume a collective conception where everyone, or some particular group of people, owned the ought claim. If we wanted, we could remove this ambiguity by explicitly mentioning the owner of the resulting obligation.[81]

Global consequentialist criterion of role-based ought

s ought that *x* be in role *R* *iff* *x* can be in role *R*;
for each other *f* that can be in role *R*, *x* is better than it in *R*;
and it isn't better to have nothing in role *R*

Here the right hand side of this formula is the same as for the criterion of rightness, but the conception of possibility involved is assumed to be tied to *s* (the owner of the ought claim).

3.3 Proposition-based global consequentialism

3.3.1 Possibilist global consequentialism

So far we have looked at the global consequentialist theories of Parfit, Railton, Kagan, and Pettit and Smith. These theories all assess focal points by considering the focal point in isolation and then looking at how changes to that focal point would affect the outcome. In contrast to this, Roger Crisp and Fred Feldman have theories in which the primary focus is on individual lives an agent could lead. Focal points are then assessed by looking at what would occur were the best life to be realized by the agent in question.

In his 1992 paper, 'Utilitarianism and the Life of Virtue', Crisp provided an argument for leading a life of virtue on utilitarian grounds.[82] To do so, he formulated a theory of utilitarianism that does not give primacy to acts:

[81] For more about 'ownership' of obligations, see the Appendix.

[82] Crisp (1992).

Biographical Utilitarianism

Any individual ought to live in such a way that the total amount of utility in the history of the world is brought as close as possible to the maximum.[83]

He then applied this principle to focal points such as acts, motives, decision procedures and character traits:

> '…according to Biographical Utilitarianism, the patterns of motivation a person ought to have are those that are the 'most useful', that is, those that make the history of the world as good as possible.'[84]

> 'And living well, according to Biographical Utilitarianism, will involve performing those acts, having those motives, following those rules and possessing that character or those traits which will maximise utility over all.'[85]

Crisp noted that by assessing such focal points in terms of what would be the case in the best available life, there is less room for conflict between, say, the motivation I ought to have and the acts I ought to perform, for these are determined in such a way that there must be a life in which I both have the motivation and perform the acts.[86]

While Crisp did not go into much technical detail about his theory, Feldman has developed a highly detailed theory along very similar lines. I shall thus focus on his account of global consequentialism for the remainder of this section, and I presume that much of what follows will also be compatible with Crisp's Biographical Utilitarianism.

In 1975 Fred Feldman wrote a paper called 'World Utilitarianism' in which he advocated a new form of consequentialism that focused on the possible lives an agent could lead. His aim was to avoid technical problems for the contemporary formulations of act-consequentialism and he used his theory to assess acts in a novel, and consistent, manner. In his 1986 book, *Doing the best we can*, Feldman kept the basic structure of world consequentialism, but extended it to assess arbitrary propositions rather than just acts. Then, in his 1993 paper 'On the consistency of act- and motive-utilitarianism: a reply to Robert Adams', he explicitly showed how his theory could be used as a form of global consequentialism (which he called 'wide consequentialism').

World consequentialism starts with the following theory of the possibility involved in 'ought' implies 'can':

---

[83] Crisp (1992), p. 142.

[84] Crisp (1992), p. 144.

[85] Crisp (1992), p. 146.

[86] Note that there may still be *some* room for conflict, since the best pattern of motivation may *in fact* lead me to perform non-optimal acts, even though I *could* have combined it with the optimal acts. I shall return to this idea in section 6.1.1.

> '…at every moment of moral choice, there are many possible worlds *accessible* to the moral agent. Consider some infant at the very moment when she begins to exist. There are various possible ways in which she might live out her life. For each of these total histories, there is a possible world — the world that would exist if she were to live out her life in that way. Call these worlds the "original stock" for the person. As time goes by, the person filters out world after world. By acting in this way rather than that, she bypasses, or rules out, many possible worlds. Once bypassed, a given world is never again accessible to a person. If we select a person, *s*, and a time, *t*, at which *s* exists, there will be a not-yet-ruled-out subset of *s*'s original stock. Each member of this subset is a possible world that is, in an important sense, still possible for *s*. I say that each such world is "accessible" to *s* at *t*.'[87]

To this Feldman adds a definition of a 'best':

> A world *w* is a *best* for *s* at *t* *iff* *w* is accessible to *s* at *t* and no better world is accessible to *s* at *t*

He then defines the fundamental principle of world consequentialism:[88]

> *World-consequentialist criterion for obligatory propositions*
>
> As of *t*, *s* morally ought to see to the occurrence of *p* *iff* *p* is true in all the 'bests' for *s* at *t*

In this formulation, *p* is an arbitrary proposition. For example, *p* might be the proposition expressed by 'Maria donates the money' or that expressed by 'Oxford has a temperate climate in the 23rd Century'. Thus if it is true that Maria donates the money in all the 'bests' accessible to Maria at time *t*, then as of *t*, Maria ought to see to the occurrence of Maria donating the money (i.e. Maria ought to donate the money). If it is true that Oxford has a temperate climate in the 23rd Century in all the 'bests' accessible to Maria on 1/1/2009, then Maria ought to see to the occurrence of Oxford having a temperate climate in the 23rd Century as of 1/1/2009.

There are a number of interesting aspects of Feldman's theory. For one, it is a tensed theory of moral obligation. The phrase 'as of *t*' is saying that the judgment is being made from the perspective of a time *t*, where different results can be reached from the perspective of different times. For example, an act being performed at a particular time t_a might be obligatory when judged from the perspective of time t_1 and impermissible when judged from the perspective of time t_2. This may seem odd, but it can be a desirable feature. An example shows why.

---

[87] Feldman (1993), p. 208.

[88] We use here the principle as outlined in Feldman (1993), p. 209. It seems likely that Feldman still prefers his more careful statement of the principle from *Doing the best we can* which traded elegance in formulation for the ability to deal successfully with cases in which there an infinite succession of accessible better worlds and no best ones. The reader should feel free to substitute in that version if so desired — nothing will turn on it here.

The follow-up

> Jacinta is suffering from a deadly medical condition and has come to see her doctor, Valerie. There are two ways to treat the condition. One is to give a dose of drug *A* now and another dose of drug *A* in a follow-up procedure in a week's time. The other is to give a dose of drug *B* now and another dose of drug *B* in a follow-up procedure in a week's time. Giving the two doses of drug *A* is slightly better than the two doses of drug *B* because it has fewer side effects. Mixing the drugs, or not giving the drugs at all will lead to Jacinta's death.

As of Jacinta's arrival at the hospital, Valerie has it open to her to provide any combination of the two drugs (or no drug at all). In the best possibility, Valerie provides drug *A* both times. Thus on Feldman's view, and according to common sense, she ought to administer drug *A* now and she ought to administer drug *A* at the follow-up appointment. But suppose that Valerie actually administers drug *B* in the first case. After this, it appears false that she ought to administer drug *A* in the follow-up. We can overcome this problem with a theory of tensed rightness such as Feldman's. At the time of the first meeting, Valerie ought to administer drug *A* at that meeting and ought to administer drug *A* in the follow-up. Since she actually administers drug *B* at the first meeting, at all times after that, she administers drug *B* in the best worlds open to her and thus *as of that time*, she ought to administer drug *B* in the follow-up.

The tensed nature of Feldman's theory can thus be seen as a virtue. Indeed a major reason for the lack of tensed rightness in other theories (such as traditional act-consequentialism) is that we typically focus on situations where the judgment about rightness is made at the same time as the act is performed. However, it is clear that we can ask about situations where this is not the case (especially when assessing other focal points), and a tensed theory plausibly resolves the ensuing issues.[89]

There is a strange quirk of Feldman's theory which is related to its tensed account of obligation. Any proposition becomes obligatory once it can no longer be made false, for it is then guaranteed to be true in all accessible worlds and thus all 'bests'. For example, suppose Michael murders someone at time *t*. As of any time after *t*, the proposition expressed by 'Michael murdered someone at *t*' will be true in all worlds accessible to Michael and as of any time after *t*, Michael *ought* to see to the occurrence of Michael murdering someone at *t*. In short, anything that can't be avoided by *s* at *t* is obligatory to *s* at *t*.

This is not obviously false, and Feldman accepts this consequence of his theory. However, we may think that just as 'ought' implies 'can', so 'ought' implies 'can fail to'. Feldman caters to this possibility by showing how a small modification can be made to his theory to remove these troubling obligations and adhere to this stronger principle. We can define the concept of a proposition being within someone's power as follows:

---

[89] Note that the need for a tensed theory is linked to the possibilist nature of Feldman's theory. Actualists can simply say that since Valerie *will* administer drug *B* first, she ought to administer drug *B* in the follow-up — even when judged from a time before the first drug is given.

p is within *s*'s power at *t* *iff* *p* is true in at least one world accessible to *s* at *t*, and
p is false in at least one world accessible to *s* at *t*

We can then add a clause requiring that a proposition must be within the agent's power:

World-consequentialist criterion for obligatory propositions (revised)

As of *t*, *s* morally ought to see to the occurrence of *p* *iff* *p* is true in all the 'bests' for *s* at *t*, and *p* is within *s*'s power at *t*

Since this revised formulation deals with the counterintuitive oughts regarding past events and has no real downside, we shall accept the revision from here on.

World consequentialism is a form of global consequentialism. It assesses acts, motives, climates, eye-colours (and everything else) in terms of the goodness of the outcomes in which they occur. In Feldman's words:

> '[World consequentialism] may be taken as a global theory of normative evaluation. It tells us what actions we should perform, what motives we should have, what traits of character we should develop, etc.'[90]

World consequentialism is also the most developed variety of global consequentialism so far. It has a consistent set of answers to each of the concerns that were discussed earlier in this chapter. Focal points (acts, motives etc.) are each assessed in the context of a given role, and this is achieved via propositions that combine the two.[91] For example we assess a given set of motives in the role of being had by Jane during May, using the proposition that *Jane has the motive-set during May*. Since propositions are very fine-grained, the theory can express a great deal. The theory also benefits from the fact that propositions are so familiar and well-studied. For example, we do not need to provide a detailed theory of action to assess acts, a detailed theory of motivation to assess motives and so forth.

World consequentialism is a variety of normative global consequentialism and provides an account of the possibility involved in 'ought' implies 'can' (or at least a sketch of the appropriate possibility). Interestingly, it is a deeply normative theory in that it does not lend itself to a single obvious way of merely *evaluating* propositions.[92]

Given the long history of world consequentialism and its degree of sophistication, it is interesting that it has been ignored in recent discussion of global consequentialism. One reason for this might be that its approach is quite different to the four varieties discussed in

---

[90] Feldman (1993), p. 212.

[91] Technically one could say that Feldman's theory does not assess acts at all, only propositions. Some philosophers may take issue with this lack of assessment of the acts themselves, but I personally don't see it as problematic since there is an obvious correspondence between judgments of acts being performed and judgments of propositions that the acts are performed.

[92] One could perhaps evaluate a proposition in terms of the goodness of the best world in which it occurs, but this would have to be a concept of the goodness of the proposition *relative to an agent s* and we typically think of the goodness of something as being agent-independent.

section 3.2. Whereas Parfit *et al* sought to assess an act (or climate, or set of motives) in comparison to other possible acts (or climates, or sets of motives), Feldman compares a world to other possible worlds and then looks at the acts (or climates, or sets of motives) that occur within these worlds.

This difference in strategies leads to an important difference in the type of theory. There is a distinction in the consequentialist literature between *actualism* and *possibilism*.[93] Actualism is the theory that we should assess things in terms of what would follow from them.[94] Possibilism is the theory that we should assess them in terms of whether they are part of the best possible course of events. Parfit *et al* have theories that are compatible with both actualism and possibilism. They could define the best motives as those which would in fact have the best results (focusing on the outcome where everyone acts as they actually will), or they could define the best motives as those which are required in order to reach the best possible results (focusing on the outcome where the agent in question chooses the best life available to her). In contrast, Feldman's and Crisp's theories are possibilist by their very nature, for their prescriptions are based solely around what happens in the best worlds accessible to the agent.

There are a number of ways in which the actualist conception of obligation better fits our intuitions. For example, when we advise our friends and tell them what they ought to do, we do better if we take into account the consequences of their acting in the way that they would in fact act, rather than assuming that they will produce the optimal outcome associated with their action. More tellingly, suppose you are considering a choice of your own where option *A* could potentially lead to a tiny amount of good, but only if you followed through and were strong enough of will to avoid many pitfalls. If you were weak of will even once, it would lead to a great calamity. In contrast option *B* is safe and boring, leading to no additional good or harm. Let us also assume that you are quite weak of will and that, while you *could* avoid all the temptations on the first course of action, you in fact won't and you strongly suspect this. It seems that it is best in such a case to settle for option *B*. However, possibilist theories like Feldman's say that you ought to choose option *A* because the best accessible outcomes involve it. This obligation sounds acceptable in theoretical terms, but it would cause great problems if it were action guiding.

This is not a fatal objection to possibilism, but shows how it clashes with certain consequentialist intuitions and thus why many consequentialists might seek an actualist theory of global consequentialism. It is for the above reasons that the role-based theory presented in section 3.2.4 was cast in actualist terms.[95] To conclude this chapter, I shall construct an actualist theory that combines some of the virtues of world consequentialism with an underlying intuition more in keeping with the approaches of Parfit *et al*.

---

[93] A good introduction is found in Jackson and Pargetter (1986).

[94] Since this is sometimes an assessment of the actual and sometimes of the counterfactual, it seems that 'subjunctivism' would be a better name than 'actualism'.

[95] It is quite simple to turn that theory into a possibilist one. We just need to change the definition of the goodness of *x* being in role *R* to look at the best world open to the agent in which *x* is in *R*. The

However, before I do so, I wish to point out an interesting possibility. While actualism and possibilism are typically seen as competitors for the 'true' account of ought, one could instead see them as two compatible senses of ought. Regarding the case above, a pluralist about actualism and possibilism would say that you ought *in the possibilist sense* take option *A* and ought *in the actualist sense* take option *B*. These senses would cut across the objective, evidence-relative and belief-relative senses, giving six senses in total. I do not wish to commit to such a pluralist view, but merely point out that it is perhaps an attractive way to sidestep the dispute and recognise that both intuitions have their place.

3.3.2 An actualist account

Following Feldman, we shall use propositions as the objects of our assessments. However, we shall have to change the method by which we assess propositions. We can begin in much the same way as for the role-based account:

Global consequentialist formula for objective goodness of a proposition

the objective goodness of *p* = the goodness of the outcome that would occur were *p* true

Global consequentialist formula for belief-relative goodness of a proposition

the belief-relative goodness of *p* = the expected goodness of *p* being true, where the probabilities are given by the agent's degrees of belief

As in the role-based approach, we can't simply use the above formulae to derive commonsense judgments about goodness due to the contextual nature of such judgments. However, we can once again use it to develop formulae for betterness and degrees of betterness of propositions:

Global consequentialist criterion of betterness for propositions

p is better than *q* *iff* the goodness of *p* is greater than the goodness of *q*

Global consequentialist formula for degrees of betterness for propositions

p is better than *q* to the degree to which the goodness of *p* is greater than the goodness of *q*

Thus, following actualist act-consequentialism, we would conclude that *Maria donates the money* is better than *Maria spends the money on chocolate* if and only if the outcome that would result from Maria donating the money is better than that which would result from Maria spending the money on chocolate.

These formulae have much in common with our earlier formulae of betterness relative to a role, and like those formulae, they must be restricted to cases where *p* and *q* are alternatives to each other. As before, we can either use the more familiar concept of hard alternatives (which cannot both be true) or the more expressive concept of soft alternatives (were one

---

disadvantage would be that judgments about betterness would have to be relative to an agent.

true, the other would be false).

Note also that these formulae for the betterness of propositions are even more expressive than the earlier formulae for betterness within roles. This is because the two propositions can refer to different roles. Consider the following example:

> *Clare and her child*
>
> Clare could either give her child some benefit, or give much greater benefits to some unfortunate stranger. Because she loves her child, she benefits him rather than the stranger.

This is a case made famous by Parfit in his discussion of 'blameless wrongdoing'.[96] It shows how having the best available pattern of motivation might lead one to perform acts which are wrong. Using the global consequentialist criterion for betterness of propositions, we can directly assess whether *Clare has this pattern of motivation* is better than *Clare benefits the stranger*.[97]

From here, the natural move is to define obligations in terms of the 'best possible'. However, this was much easier for the role-based approach as we were looking for the best member of a given class. When it comes to propositions, we can only move from better to best in the context of a set of alternatives. Rather than attempt to specify the appropriate alternative set, let us begin by looking merely at an agent's obligations relative to a given set of alternatives.

> *Global consequentialist criterion of ought-rather-than for propositions*
>
> As of t, s ought that p rather than $q_1 \ldots q_n$ *iff* p, $q_1 \ldots q_n$ are each within s's power at t, and p is better than each of $q_1 \ldots q_n$

This criterion bears many similarities to world consequentialism: it uses a tensed ought, it separates the owner of the obligation from the contents of the obligation and it specifies 'can' in terms of the worlds accessible to the agent (via the former definition of 'within one's power'). However, unlike world consequentialism, this is an actualist theory. Consider the following case:

> *A predictable slump*
>
> Brian is deciding whether to stay home or to go to the pub. He hopes to work if he stays home, but will instead slump in front of the television. The outcome of staying home and working is the best, followed by going to the pub, followed by staying home and slumping.

The criterion claims: 'Brian ought that *Brian goes to the pub* rather than *Brian stays home*', because it is within Brian's power that he goes to the pub and the outcome that would result from him going to the pub is better than the outcome that would result from him staying home (and consequently slumping in front of the television). Interestingly, the criterion also claims that 'Brian ought that *Brian stays home and works* rather than *Brian goes to the pub*'.

---

[96] Parfit (1984), pp. 32–5.

[97] It is also possible to extend the role-based approach to make cross-role comparisons like this one.

Contrary to appearances, there is no conflict between these ought claims because they are made relative to different sets of alternatives.[98]

Before we provide the last piece in this global consequentialist theory, let us consider just what is missing. Suppose that we want to know whether Maria ought that Maria donates her money. While our theory does not answer this, it does answer a lot of very similar questions. In particular, for each set of alternatives S, it tells us whether Maria ought that Maria donates her money rather than any alternative in S. Is this enough? In one sense it seems that it is. For suppose that the relevant set of alternatives is some set of actions S^*. Now one of the things that the theory already tells us is whether or not Maria ought that Maria donates the money rather than any alternative in S^*. It just does not do so under a special description where the alternative set is not explicit. If we knew which were the relevant alternatives, then we would not need to know any more, and arguably answering this question of relevance is not the responsibility of the moral theory, but of the person posing the question.

I thus propose the following analysis of ought claims where no alternatives are specified:

> 's ought that p' ambiguously stands for some claim of the form: s ought that p rather than $q_1 \ldots q_n$. There is no single correct interpretation as it depends on speaker and context.

Thus, on this account we could say both that 'Brian ought to stay home and work' and that 'Brian ought to go to the pub' without any explicit conflict, for the obvious interpretation is that these prescriptions are made from different implicit alternative sets.

3.4 Strengths and weakness of global consequentialism

3.4.1 The merits of global consequentialism

We have explored the possibilities of global consequentialism at length and are now in a position to reflect upon its merits.

The major advantage is that global consequentialism is much more expressive than act- or rule-consequentialism. Instead of being restricted to judgments about one or two focal points, it can assess every possible focal point, allowing it to capture a significantly wider range of everyday moral expression. Specifically, its ability to assess decision procedures and motives opens the door to being able to correctly formulate and answer the questions of decision making and motivation that were outlined in Chapter 1.

As we saw in Chapter 2, act-consequentialism can be used to obliquely make moral assessments of many focal points by asking whether we should perform acts which would lead, say, to the internalization of a certain decision procedure or the existence of a certain climate in the year 2300. However, these are not the same as direct assessments, for we can imagine cases in which it would be better if I currently followed a certain decision procedure or had a certain pattern of motivation, but in which some catastrophe would occur if I acted

[98] See Jackson and Pargetter (1986).

to make this the case. This strategy of oblique assessment can thus imitate the direct assessment of other focal points, but the imitation is not always perfect. If translating ordinary statements which appear to assess our motives or decision procedures into oblique and sometimes inaccurate statements about associated acts was the only possibility open to consequentialists, then it might not appear so bad. However, when global consequentialism has been made explicit, it seems like a much better option.

Chapter 2 also showed that rule-consequentialism can be used to assess both acts and decision procedures, but that these assessments are indirect in the case of acts, and that such indirection is in conflict with the fundamental consequentialist intuition. In contrast, global consequentialism assesses acts and decision procedures directly in terms of their consequences and thus stays true to this consequentialist tenet, never preferring the lesser good to the greater.

The two global consequentialist theories that I developed are also more expressive in a second way. Act-consequentialism only considers the effects of the agent in question performing the act and rule-consequentialism only considers the effects of everyone (or almost everyone) following the set of rules. In contrast, global consequentialism can make both types of judgment. It can determine whether it is better that *Jill lies in the current situation* or that *Jill tells the truth in the current situation*, and also whether it is better that *everyone lies when they think it is worth it* or that *everyone always tells the truth*. Since people regularly ask such questions, this is another notable advantage in terms of expressivity. It may have an air of paradox, but that is only because we are familiar with rule-consequentialism and Kantianism which make judgments in the individual case based on the outcomes of the universal case. In contrast, global consequentialism can make both judgments and keep them separate. If you want to know what it is best for everyone to do, the theory gives you one answer; if you want to know what it is best for Jill to do, it gives you another. If Jill can only control what she does, then she would probably be more interested in this latter answer, but she may still have an academic interest in the former and global consequentialism can provide this.

The other notable advantage of global consequentialism is that it is a simpler theory than act- or rule-consequentialism. Although this chapter has been quite complex, that is because it has covered many different versions of global consequentialism and several issues which affect all types of consequentialism. Global consequentialism is simpler than act- or rule-consequentialism because it does not have to carefully pick out the exact sense of act or rule that is relevant for moral assessment, nor the exact role in which they are to be assessed.

For act-consequentialists, it is essential to know how fine-grained the account of acts is supposed to be. There are also questions about whether the composition of two acts is itself an act, and the same goes for conjunction, disjunction and negation. Rule-consequentialists debate whether we are to look at *compliance* with rules, or *being guided* by rules, or at *internalization* of rules. They also debate whether we are to consider the rules being followed by *everyone*, or by *the overwhelming majority* and just how large such a majority needs to be.

In contrast, this complexity does not need to be built into global consequentialism: it comes in through the questions we ask. Suppose we label the competing accounts of acts as 'act_1',

'act_2', and so on. Global consequentialism says that the right act_1 for someone to perform is the act_1 which leads to the best outcome, the right act_2 is the act_2 which leads to the best outcome etc. It can tell us which is the right set of rules for everyone to comply with and, simultaneously, which is the right set of rules for at least ninety percent of people to internalize. Global consequentialism is thus a highly pluralist theory. As well as allowing assessment of many different types of focal point in many different roles, it provides simultaneous assessment of every nuanced account of each focal point or role. For global consequentialists, the substantive question of which account of acts or rules should feature in the theory becomes a very different type of question: of the myriad pronouncements of global consequentialism, which ones are we interested in? Indeed, this question forms the basis of the next two chapters, with a focus on decision procedures in Chapter 4 and motivation in Chapter 5.

It may be thought that act-consequentialism is simpler than global consequentialism precisely because it makes fewer claims. However, this would be to mistake a deficiency for a virtue. Consider the theory of Tuesday-act-consequentialism that makes act-consequentialist judgments of the rightness of acts performed on Tuesdays and leaves all other acts unassessed. Though this theory in turn makes fewer judgments than act-consequentialism, it is worse on two counts: it is less expressive, remaining silent on many acts that we would like to assess, and it is more complex, inserting an arbitrary complication into its criterion. Similarly, act-consequentialism's lack of judgments about non-act focal points means that it is both less expressive than global consequentialism and that it has unnecessary complexity with its detailed prescriptions about exactly which focal point it will assess.

3.4.2 Objections to global consequentialism

Many objections to act-consequentialism are also objections to global consequentialism since it makes exactly the same claims regarding acts. No doubt some rule-consequentialists, deontologists, and virtue ethicists will reject it for the same reasons that they reject act-consequentialism. In Chapter 7, I will do my best to counter these objections by showing that global consequentialism can go some way to bridging the gap between consequentialism, deontology and virtue ethics. Although it makes all the claims that act-consequentialism makes, global consequentialism makes them from within a framework that takes decision making and motivation seriously and this can make a large difference.

However, I shall not attempt to respond in detail to the individual objections that could be raised at global consequentialism in virtue of its similarity to act-consequentialism. While formulating the strongest responses to these criticisms would be a very worthy project, it would require its own book-length treatment and I am unable to pursue this here. Instead, I shall restrict my attention to criticisms that are of particular relevance to global consequentialism itself.

The most prominent of these concerns the possibility of conflict between the various ascriptions of rightness. For example, Hooker suggests that there might be such a conflict between the decision procedure that it is right for an agent to follow and the acts that it is

right for her to perform.[99] In particular, he suggests that following the right decision procedure will sometimes involve the agent performing an act that is wrong. This does not appear to be a *logical* inconsistency, but it certainly hints at a most unintuitive aspect of global consequentialism and Bart Streumer has suggested that it conflicts with the rule that 'ought' implies 'can'.[100]

There are also a number of other objections. For example, Rawls and Williams have each suggested that an approach like global consequentialism is no longer distinctively consequentialist and that in allowing the best decision procedures and motives to come apart from naïve calculation and impartial benevolence, consequentialism 'ushers itself from the scene'.[101] In addition, Hooker has criticized global consequentialism for assessing the holding of attitudes such as blame and guilt in terms of their consequences, arguing that these attitudes play a special role in ethics and the conditions for when they should be held cannot be given in consequentialist terms.

These are important objections and to do them justice will require an entire chapter. I therefore dedicate Chapter 6 to responding to them. I am not entirely happy about delaying this response for so long, but the objections can be better understood after we have looked in depth at the nature of decision procedures and motivations, and also at the roles in which they can be assessed. I shall thus move on to these matters now, temporarily setting aside these objections. Then, in Chapter 6, I shall discuss them at length and explain how they can all be satisfactorily resolved.[102]

Chapter summary

In this chapter I examined six pre-existing theories of global consequentialism and their different approaches to assessing everything. I paid particular attention to: the roles in which things are to be assessed, the question of whether normative terms are to be applied to things other than acts, and whether the theories were actualist or possibilist.

In light of this analysis, I developed two new theories of global consequentialism: one based around explicit roles and the other around propositions. Either of these theories can provide the assessment of decision procedures and motives that we are searching for. In the following chapters, I shall use the role-based theory of global consequentialism from section 3.2.4, and will freely use normative terms such as 'right', 'ought' and 'should' to assess things other than acts. However, if you prefer either of the proposition-based theories, or to restrict assessment to evaluative terms, then I believe I have provided enough detail here to allow you to

---

[99] Hooker (2004), § 5.

[100] See Streumer (2003) and Streumer (2005).

[101] See Rawls (1971), p. 182 and Williams (1973), p. 134. The exact quotation from Williams is that 'utilitarianism's fate is to usher itself from the scene'.

[102] If you feel that one of these objections is particularly damaging to my project and are unsatisfied by this delay, then by all means skip straight to the appropriate section of Chapter 6 to see whether I resolve it to your satisfaction.

translate my claims in subsequent chapters into these alternative forms of global consequentialism.

In the next two chapters, I shall show how global consequentialism can be used to answer the question of decision making and the question of motivation. This will require an analysis of the different ways in which we could make these questions precise, looking for the most apt account of decision procedures and motivations, as well as the roles in which we are to assess them. Only when this is done will I be able to properly address the criticisms that people might have to the global consequentialist approach, and I devote Chapter 6 to this topic. Once the objections have been addressed, I shall briefly explore what this extension of consequentialism means for the relationship between consequentialism, deontology and virtue ethics.

Chapter 4

Interpreting the question of decision making

So far, I have left the question of decision making in a rough form: 'How should I decide what to do?' This was sufficient for the purposes of Chapter 2, in which I showed that act- and rule-consequentialism could not adequately answer it, and for Chapter 3, in which I developed the theoretical machinery needed to assess everything according to its outcome. However, in order to usefully apply global consequentialism to the question, we need to be more clear about what it is asking.

The question of decision making is asking for a way of deciding what to do — a decision procedure. To apply global consequentialism to this question, we need to phrase it in a way that assesses a decision procedure in a particular role. To begin with, we can phrase this quite roughly as follows:

> *Q*: 'Which decision procedure should I follow?'

From this phrasing, we can see that the object of assessment is a decision procedure, and the role in which it is to be assessed is that of being followed by the agent. However, this question is currently ambiguous. There are many ways in which we could flesh out what we mean by 'decision procedures' and 'following'. For every way in which these could be fleshed out, we get a different sharpening of *Q*.

Global consequentialism answers all such interpretations of *Q*. However, some of these interpretations will be more interesting than others. The purpose of this chapter is not to legislate a single 'correct' interpretation, but to explore the different interpretations and look at their properties. As we shall see, some of them are distinctly more relevant to ethical life than others.

In section 4.1, I provide an overview of what consequentialist authors have said about the nature of decision procedures. In section 4.2, I develop these thoughts and flesh out the concept of a decision procedure. In section 4.3, I do the same for the concept of *following* a decision procedure and reach some firm conclusions. Finally, in section 4.4 I explore a number of additional features of the question of decision making, such as whether it is better interpreted as being concerned with objective rightness or with subjective rightness.

4.1 Prior approaches

While there is widespread agreement among consequentialists that agents should not make all of their decisions via naïve calculation, their alternative analyses of the issue are disparate and often lacking in specifics. In this section, I survey these prior approaches.

4.1.1 Habits and predispositions

In *Utilitarianism: for and against*, J. J. C. Smart discusses the problems with naïve calculation:

> 'Consider a case in which a man sees that his wife is tired, and simply from a spontaneous feeling of affection for her he offers to wash the dishes. Does utilitarianism imply that he should have stopped to calculate the various consequences of his different possible courses of action? Certainly not. This would make married life a misery and the utilitarian knows very well as a rule of thumb that on occasions of this sort it is best to act spontaneously and without calculation.'[103]

He points to the fact that we do indeed 'habituate ourselves to behave in accordance with certain rules, such as to keep promises'[104], and explains how these 'rules-of-thumb' free us from perpetual calculation. However, he regards these rules-of-thumb, not as any kind of decision making procedure, but as some lesser cognitive process: a mere habit or disposition. Thus Smart writes:

> 'He acts in accordance with rules, in short, when there is no time to think, and since he does not think, the actions which he does habitually are not the outcome of moral thinking. When he has to think what to do, then there is a question of deliberation or choice, and it is precisely for such situations that the utilitarian criterion is intended.'[105]

Smart thus separates our cognitive, action-producing powers into two classes: on the one hand there are decisions or choices, which proceed by deliberation, and on the other hand there are instinctive, habitual responses via rules-of-thumb. He proposes that whenever we truly deliberate, we should use some kind of naïve calculation, but that in many cases, it is better that we act on mere rules-of-thumb, bypassing deliberation altogether. He also provides a kind of indirect normative approval of rules-of-thumb by pointing out that since rules-of-thumb can be so beneficial, naïve calculation recommends inculcating them within oneself and others.

Pettit and Brennan offer a very similar distinction in their paper *Restrictive Consequentialism*. Where Smart talks of habit, they talk of mental states of a certain type which they name *predispositions*:

---

[103] Smart (1973), p. 44.

[104] Smart (1973), p. 42.

[105] Smart (1973), pp. 42–3.

> 'They are states whose manifestation in action means that the action is not chosen on a fully calculative or deliberative basis.'[106]

As in Smart's account, the alternative to naïve calculation is considered to be something that can 'pre-empt certain decisions', rather than forming a natural part of decision making.[107] However, on this account the predispositions aren't claimed to be completely distinct from deliberation, but are just not *fully* deliberative.

This approach of Smart, Pettit and Brennan is very interesting, for it is an attempt to avoid the problems of naïve calculation without attempting to provide a consequentialist evaluation of decision procedures in an analogous way to acts. Instead, it gives naïve calculation the privileged place of being the recommended consequentialist decision procedure, but limits the damage by restricting it to a subset of all cognition about which action to perform. This approach is therefore not an example of the 'standard consequentialist response' that I am attempting to flesh out in this essay. This might not be a problem for the approach if it could be made to work, but as I shall show, it fails precisely because it does not assess the outcome of using a certain decision procedure.

I agree with Smart, Pettit and Brennan that the selection of actions via unthinking rules-of-thumb or predispositions can be very useful. However, this does not go far enough, for deliberation via complex decision procedures other than naïve calculation can also be very useful. To see why, consider the following example due to Railton.[108]

> *The demon of decision making*
>
> The world is held to ransom by an all-knowing all-powerful demon. This demon examines the way in which each person makes their decisions and causes horrific pain and suffering whenever someone fails to make decisions according to a particular Kantian decision procedure of attempting to determine the maxims under which they act and whether or not these are universalisable.

In such a case it would be best for all of us to begin using the Kantian decision procedure described above.[109] This would clearly involve actual deliberation and more than just habitual responses. We would thus be in a situation where it would be optimific for us to *decide* in a way other than naïve calculation. Indeed, we need not go so far as demons. Smart, Pettit and Brennan admit that rules-of-thumb or predispositions can allow us to produce better actions overall than naïve calculation for such reasons as time efficiency and personal biases. Surely though, such reasons also bear upon situations that cannot be addressed

---

[106] Pettit and Brennan (1986), p. 440.

[107] Pettit and Brennan (1986), p. 440.

[108] Railton (1984), p. 155. Note that while Railton uses this to argue forcefully for our using the optimific decision procedure (or 'mode of decision making'), he says very little about what a decision procedure is or in which role we are to judge it (compliance, acceptance etc.).

[109] Perhaps this would be best achieved by attempting to convince everyone that consequentialism is actually false (making consequentialism a *self-effacing* theory), perhaps not. It does not matter to the present discussion.

merely by a non-deliberative instinctive rule. A certain form of deliberation may well do better than both naïve calculation and any rule-of-thumb: we cannot rule this out *a priori*.

4.1.2 Rules and principles within decision making

When Mill confronts the objection that consequentialism cannot be practically applied, he states:

> '…mankind must by this time have acquired positive beliefs as to the effects of some actions on their happiness; and the beliefs which have thus come down are the rules of morality for the multitude, and for the philosopher until he has succeeded in finding better. … The corollaries from the principle of utility, like the precepts of every practical art, admit of indefinite improvement, and, in a progressive state of the human mind, their improvement is perpetually going on. But to consider the rules of morality as improvable, is one thing; to pass over the intermediate generalisations entirely, and endeavour to test each individual action directly by the first principle, is another.'[110]

While we may well question Mill's claim that we have our rules of common-sense morality *because* they are believed to be conducive to universal happiness, it is difficult to deny that they indeed *are* conducive to universal happiness. Since such rules give us a way of acting which promotes the good (in accordance with the 'first principle'), Mill advocates their use and names them 'secondary principles'. In this way, the secondary principles are merely derivative of the first principle, but (contrary to Smart) are not claimed to be mere habits of mind. Mill does not explain precisely how we are supposed to use the secondary principles, but he does liken their use to that of landmarks which guide us on a journey.[111] From this and their being named as 'principles', it would appear that they are to be appealed to in deliberation and not restricted to mere automatic action responses.

G. E. Moore also claims that we should make decisions according to a set of rules which are roughly modelled on common-sense morality.[112] He explains why following these rules will generally lead to the best outcome and then that we can on no particular occasion be justified in believing that we are in one of the exceptional cases in which we should break the rule.[113] This extreme pessimism regarding our knowledge of exceptional cases leads to a very strong view on which the 'principles by which the individual should decide' are to consist purely in the rules of common-sense morality, and leave no room for any appeal to the goodness of outcomes. However, this is an empirical assumption on Moore's part (and a dubious one at that) so we should leave open the analytical possibility of the best principles involving some form of calculation.

---

[110] Mill (1861), Ch. 2, pp. 275–6.

[111] Mill (1861), Ch. 2, p. 276.

[112] Moore (1903), pp. 152–67.

[113] Moore (1903), pp. 162–4.

R. M. Hare provides a very sophisticated account of moral decision making via principles.[114] He divides these principles into two levels: the *intuitive level* and the *critical level*:

> 'Intuitive principles are for use in practical moral thinking, especially under conditions of stress. They have to be general enough to be 'of ready application in the emergency', but are not to be confused with rules of thumb (whose breach excites no compunction). Critical principles are what would be arrived at by leisured moral thought in completely adequate knowledge of the facts, as the right answer in a specific case.'[115]

Hare thus fleshes out Mill's distinction between his first and the secondary principles. Like Mill, Hare points out that people actually do use intuitive principles, such as 'Do not commit adultery' or 'Protect your children' and that, while they are not of primary theoretical importance, the use of such principles leads to a significant amount of good. The critical principles are then used to select new intuitive principles to indoctrinate in others, to help decide between conflicts of intuitive principles and to modify our own intuitive principles when we become aware of a systematic failing in them. This explanation in terms of the two levels of moral thinking fits well with our pre-theoretic moral reasoning and also gives a fairly explicit account of assessing sets of intuitive principles in terms of the benefits of possessing them:

> '…the object being to have those intuitive principles whose cultivation and general acceptance will lead to actions in accord with the best critical principles in most situations that are actually encountered.'[116]

This does, however, seem a little confused. If we adopt consequentialist critical principles (as Hare believes that we should) then we must take care to note that the intuitive principles whose acceptance would lead to the best possible outcome are not necessarily those whose acceptance would accord with the critical principles in most actual situations. In particular suppose that my acts are all wrong by the act-consequentialist criterion, but are each only slightly suboptimal. This may well lead to more good than if only one of them was wrong, but it was catastrophically wrong. It thus seems best to amend Hare's claim so that intuitive principles which lead to the most good are to be preferred.[117]

---

[114] Hare (1981), Ch. 3; Hare (1989), Ch. 4 & Ch. 13.

[115] Hare (1989), p. 221.

[116] Hare (1989), p. 221.

[117] Hare presumably used his phrasing because it is applicable to both deontological and consequentialist critical principles. The problem is that it gets the consequentialist approach slightly wrong as I have shown. I do not know how it could be rephrased so as to work for both types of critical principle. This issue often comes up when trying to formulate ethical concepts broad enough to accommodate both consequentialist and deontological approaches; for example it makes it difficult to formulate a general definition of

Hare stresses the importance of judging the intuitive principles based on the good that comes from their being *accepted*, rather than their being *complied with*.[118] There are several significant advantages that come from such an approach. Firstly, as stated above, this is taken to involve some form of deep commitment to the principles as opposed to their being treated as a mere calculational convenience. This helps ensure that the agent does indeed keep to the rules in question and take them very seriously. Secondly, it allows the complexity of the principles to be taken into account, for there is a limit on how complex a set we can commit to. Moreover, rules which are quite complex or quite demanding of the agent are more difficult for the agent to correctly apply, so these defects will be taken into account when the rules are being assessed. Finally, there may be benefits of acceptance which do not manifest in the acts that are produced. Typical examples involve the expression of a person's principles and feelings in her countenance. It may be that someone's benevolence to her friends or her willingness to retaliate are seen clearly in her expressions and difficult to fake, leading to additional benefits from the acceptance of these principles even without any associated acts taking place.[119]

The most striking feature that Hare introduces to his account of consequentialist decision making is the split into two levels. He can thus account for the necessity of rules which provide for fast, reliable, unbiased decision making, and also for more considered decision making when time allows or when the situation is particularly unusual. A particular benefit of this is that it allows for the agent to reshape their intuitive principles in light of information regarding their usefulness. Let us take the agent's entire decision procedure to consist of the intuitive level, the critical level, and the rules that determine how these related to each other. We can then see that Hare's system allows for this overall decision procedure to demand that we modify parts of it in light of new information. We shall call such a decision procedure *self-modifying* (although of course it is the agent who performs the modifications at the decision procedure's request).

However, there is also a problem with Hare's two level approach. While it allows a flexible choice of principles at the lower level, the upper level is fixed. It is to be a form of consequentialist criterion of rightness against which intuitive principles and exceptional cases are judged. In other words, a form of naïve calculation aimed not at selecting acts, but at selecting rules for the intuitive level.

It seems perfectly possible, though, that we might to do better in consequentialist terms if we adhered to some form of Kantian principles at the higher level and hence subscribed to various intuitive principles which were justified by this Kantianism. For instance, this might be true in the example of the *Demon of decision making* (from section 4.1.1). Alternatively, we might do better with just a set of intuitive principles and no recourse to a higher level, or with a system in which there were three or more levels in the hierarchy. This is an empirical matter and it does not seem right to construct the theory on the empirical assumption that

---

subjective rightness.

[118] This distinction will be drawn out later in this chapter.

[119] I cover these points in more detail in section 4.3.3.

the two level structure will lead to the best outcome.[120] Instead, we would do best to conceive of Hare's two level approach as a very plausible example of a form that the best decision procedure may take.

4.2 The nature of decision procedures

Let us now attempt to put together a precise account of the nature of decision procedures. As I have (loosely) defined it, a decision procedure is a 'way of deciding what to do'. A decision involves an agent who has a certain epistemic state and a set of available acts.[121] What we are interested in is the mental process which takes the agent's epistemic state as input and produces an intention to perform a particular act.

As Smart has pointed out, there are conscious and unconscious mental processes with which we can produce acts. In this essay, I wish to focus on mental processes that can be genuinely termed 'decisions'. I thus do not wish to consider, for example, the instinctual process whereby one's hand is quickly withdrawn when it touches something very hot. In practice it can be quite difficult to tell exactly how someone arrived at a particular intention. For example, at times it is easy to tell that I am using a particular principle to determine how I will act, but at other times — when one option is obviously superior — it is difficult to tell if there is a very short period of similar reasoning or if I have decided by unconscious 'instinct'.

However, for the purposes of this essay I shall stick to a loose conception of decision making in which some form of mental deliberation occurs (for example, relevant facts are weighed against each other, or principles are applied) and accept that in some situations it may be difficult to tell whether or not this has occurred and thus whether or not a decision has been made. Thus, unlike Smart, I shall generally consider the applications of very simple rules-of-thumb to count as decisions and (as explained above) all decisions whether short or long are open to assessment on consequentialist grounds.

The way in which we decide to do something depends on more than just our broad aims and motives. There is no doubt that these affect the decision making that does occur, but they do not completely determine it. Instead, one can have a particular set of aims and motives and yet be capable of deciding what to do in various different ways. For example, I might aim to keep my promise and to help my friend, yet I could still decide what to do by calculation, or by trusting to instincts, or by appealing to various principles. These are different methods, or procedures, by which we can move from our information to action.

---

[120] Hare's reasons for building in this inflexible assumption stem from his irrealist meta-ethics. In my formulation, there are really three places for principles: the intuitive level, the critical level and then the principles which are actually true of the world. For Hare (and many others) there is no room for such a third level at which the consequentialist criterion could reside and it is thus built into the critical level.

[121] One could instead consider a set of *options* which may include more than just acts. Depending on one's view of acts, the class of options may be more general. However, the question at hand involves deciding what to *do* and we shall consider all things that the agent can decide do as acts. By this, I only mean to make a terminological and not a metaphysical assumption.

Such procedures typically involve an appeal to a set of principles, the performance of a certain calculation, or the weighing of consequences, but they need not be limited to such things. Indeed, these procedures are algorithmic in nature and can involve common algorithmic structures, such as those involved when we perform long division. For example, decision procedures could include conditional branches, manifested in patterns of thinking such as: 'Does condition *x* hold? If it does, then carry on the decision making in such and such a manner otherwise, carry on like this…' They could also involve repetition of the form 'Keep doing *x* until condition *y* is satisfied.'

However, there are also features of our decision procedures which are not found in the algorithm for long division. For one thing, we can perform various intuitive steps. Even when we are trying to calculate which of two courses of action will lead to the best outcome, we do not perform any explicit multiplication of the probabilities and values, but typically just think about the major possibilities for a while and then come up with an intuitive judgement. In effect something like a calculation of expected goodness is achieved, but there is no conscious 'carrying the one' taking place. These intuitive judgements are much faster than the explicit ones would be and it is hard to imagine how a decision procedure could be practical without making significant use of them.

We also perform other mental tasks which are not found in mathematical algorithms or culinary recipes. For example, we can imagine the amount of suffering that a certain person would feel if we performed an action and compare it with some other imagined suffering. We can also try to change our emotional state in order to empathise with someone or to calm ourselves down. Indeed, our decision making processes can call for us to perform any mental activity of which we are capable.

Finally, as Hare stressed, we need not just follow certain procedures mechanically, but can attach great intuitive moral weight to them. For example, a utilitarian might have an (overrideable) rule against lying and also feel a strong sense of repugnance when they consider telling a lie. Since this additional compunction will make it easier for the utilitarian to keep to the rule when he is suffering from weakness of the will, it improves the decision procedure.

An important issue that remains is the granularity of decision procedures. For example, when I ride my bicycle, I do so via a sequence of fine-tuned movements of my body. These movements are determined by my mind's processing of information from my environment, such as what I can see and my internal sense of balance. Moreover, the bicycle, my environment and me, together form a continuous dynamical system. The movements of the bicycle affect my mind, which causes new movements in my limbs, which move the bicycle and so on. In some sense, my actions are being determined by a decision procedure. It is not a conscious one and not a discrete one, but my mind is certainly transforming information into actions.

Similarly, suppose that you are batting in a game of cricket and are ready to hit the incoming ball. You have a certain procedure for hitting the ball: if it bounces at your feet, then just block; if it comes high on your left hand side, then sweep the bat across. In this case the

situation is less dynamic, but there is still a question of how fine grained the decision procedure can be. It seems clear that a decision procedure can distinguish a block from a sweep, but what about different kinds of blocking strokes? How fine-grained an action can be considered? Is there a decision procedure that tells you to hit it in such-and-such a way — a way that happens to be the very best movement that you could make with your body?

In my view, the bicycle case requires so little conscious attention that it is probably best thought of not as a decision procedure at all, on the grounds that were given at the beginning of the section. There is some kind of information processing going on though, and even if it doesn't seem to be involved in the question 'How should I decide what to do?', there are other questions such as 'How should I process my information into action?' in which it might be quite relevant.[122]

The second example involves the question of how fine-grained individual actions can be in general and, specifically, how fine-grained they can be in decision procedures. In fact, there are two related features at play here: how fine-grained the actions can be and how fine-grained the input information can be. For example, we might have a decision procedure saying that if the ball is *precisely* 20° or more to the side, then sweep and otherwise block. It is not clear what to say about such 'decision procedures'. Even if we reject the purported bicycle riding decision procedure on the grounds of being too instinctual to count as decision making, that does not imply that we must reject this very fine-grained information and action.

I am unsure as to which level of granularity is of primary interest when we ask 'How should I decide what to do?'. On the one hand, when a decision procedure is proposed, it is done in quite a course-grained manner. This is partly for the ease of communication, but there is also an intuition that a decision procedure is sufficiently coarse-grained that several different people could all be said to follow the very same decision procedure. On the fine-grained account, this is unlikely to be possible since the slight differences in the ways that we interpret it will mean that we are actually following different (fine-grained) decision procedures.[123]

On the other hand, consider a tiny improvement to my current decision procedure (understood in a coarse-grained way). For example, it might be that I could weigh in other people's feelings very slightly more than I currently do. If we want to be able to say that I should decide in this new way in preference to the old way, we would seem to require a fine-grained account of decision procedures. I shall thus remain uncommitted on the issue of granularity, accepting that opinions will differ on which account is the more natural here.

Finally, I should say something about freedom. A decision procedure need not be a straight-jacket. We have already seen that plausible decision procedures will require their users to make certain judgments intuitively. This role for intuition could be quite broad. Thus, we could have a minimalist decision procedure which says 'Do whatever you want so long as

---

[122] A focus on this question would fit well with a fine-grained theory of action in which all bodily movements, however precise, are individual actions.

[123] See also Holly Smith's discussion of this case in Smith (1988), p. 90.

you never kill', though it is unlikely to lead to the best outcome.

Decision procedures which are extremely restrictive will be very difficult to follow, and will not be compatible with the agent having a fulfilling life for him or herself. The best decision procedures to follow will therefore include some form of freedom for the agent and this forms an analogue to the agent-centred prerogatives within deontological theories. While consequentialism strictly speaking does not have such prerogatives, it will recommend decision procedures that allow at least some freedom for the agents; where the degree of freedom will depend upon how easily the agent can do good for others compared to improving her own life, and on how great a strength of will she possesses.

4.3 'Following' a decision procedure

Having established some important aspects of the nature of decision procedures, we can now turn to the role that they are to play in framing the question of decision making. When we assess something using global consequentialism we must assess it in a specified role. We will start from the question 'What decision procedure should I follow?' and investigate several senses of the term 'follow' which each lead to a different interpretation of the question of decision making. As we shall see, these are all valid global consequentialist questions and have different answers: the right decision procedure to follow in one sense may not be the right decision procedure to follow in a second sense.

4.3.1 Compliance

One obvious way to understand *following a decision procedure* is in terms of *complying with the decision procedure*. It is with respect to compliance that the early forms of rule-consequentialism were phrased, and it is a natural approach. An agent can be said to comply with a set of rules if her actions are consistent with them. Thus, an agent who lies at any time has failed to comply with the rule 'Never lie', while an agent who never lies (even if this is merely by never having an opportunity to do so) does comply with this rule.

Let us then consider the question:

> Q_1: 'Which decision procedure should I comply with?'

There are at least two ways in which this concept of compliance can be understood when it is applied to decision procedures rather than to rules. For the purposes of this essay, I shall say that an agent complies with a decision procedure if and only if she performs the action(s) that correct application of the decision procedure would dictate. Note that there is an alternate notion that could claim the name 'compliance' here: that of following every step of the decision procedure perfectly (and thus arriving at the correct act). This second concept is a stronger one and I shall discuss it in detail under the name 'perfect execution', but for the present purposes I shall take compliance to be defined in the manner above.

A significant problem with the compliance view is that it judges a decision procedure on what it rules out, but not upon how it rules it out. Thus, decision procedures that involve time consuming calculations or are very prone to errors in their execution are not penalised for these features. In this way, many of the problems with naïve calculation that were discussed earlier are not going to be taken into account by this formulation of the question of decision making. The right decision procedure to comply with is thus probably going to be one that makes heavy use of complex rules to specify the actions to be performed.[124]

It is also worth noting that the compliance approach cannot distinguish functionally identical decision procedures: those that lead us to the same acts, even if by different methods. For any decision procedure we can always construct a modified version that dictates the same actions, but which does so in a slightly different way. There will thus always be a collection of decision procedures that are right for me to comply with (and no individual decision procedure that I ought to follow).

We have seen that the compliance view does not take sufficient regard of the complexity of decision procedures, but there *is* an important manner in which complexity is relevant on the compliance view. There are two different ways in which a decision procedure can be complex. Let us say that it is *internally complex* if it directs the user through a difficult sequence of mental steps along the way to producing an action and it is *externally complex* if it produces a sequence of actions that could *only* be produced by an internally complex decision procedure. While all externally complex decision procedures are internally complex, the reverse is not true because the internal complexity might just end up producing a simple sequence of actions which could in turn be produced by a simpler decision procedure. The compliance view ignores internal complexity, since it is not concerned with how the actions are chosen. There is no difficulty complying with a decision procedure that leads to simple actions but via highly complex routes, for the agent can arrive at the actions by other simpler routes and still be compliant.

However, suppose that a certain decision procedure is externally complex, and thus that its resulting actions can only be produced by internally complex decision procedures. In this case, the compliance account will take some account of the costs of following this decision procedure. For example, if the external complexity makes the decision procedure very slow to use, then it would be impossible to comply with it without taking a long time to make decisions and this cost would thus be factored into the assessment.

The compliance view therefore does take complexity into account to some degree, but it is not enough to make a practical difference. It does not penalise a decision procedure for being overly complex so long as there is some decision procedure that is free of these flaws and still

---

[124] Note that rule-consequentialism was also charged with a very similar type of collapse which we should take care to distinguish here. On a compliance account, rule-consequentialism could be said to collapse not into the rule of naïve calculation, but to the rule 'Do what will make the outcome best'. This cannot happen in our study of the compliance account of decision procedures because 'Do what will make the outcome best' is not (on its own) the type of thing that could guide our action and thus not a decision procedure at all.

leads to the same results. Moreover, it does not penalise a decision procedure at all for being susceptible to error of calculation or bias. Since it pays so little regard to the actual structure of decision procedures, the compliance view cannot be considered to be a satisfactory interpretation of the question of decision making.

4.3.2 Perfect execution

As discussed above, we could also assess decision procedures in terms of the good that would result when we follow all their steps perfectly. We could thus ask:

> Q_2: 'Which decision procedure should I execute perfectly?'

Unlike compliance, this condition allows us to account for the length of time that it would take to follow the decision procedure perfectly. For example, if someone is drowning in the river, it may be that it is much better to jump in and save him, than to do nothing, but better still to call out for assistance. Suppose that decision procedure D_1 recommends jumping in to help and gives this guidance very quickly, while D_2 ultimately recommends calling out for assistance, but requires ten seconds of thought to reach that conclusion. In this case it is much better if we perfectly execute D_1 than if we perfectly execute D_2, even though complying with D_2 may be better than complying with D_1. We can thus see that the decision procedure singled out by this version of the question of decision making would be a compromise between guiding you to good acts and guiding you quickly. Indeed, one could imagine cases where being guided more slowly is better or in which the good comes from some other consequence of having worked through the procedure. Importantly, the level of abstraction is no longer so high as to obscure most of the problems with naïve calculation.

It is not always clear what the agent must do in order to perfectly execute a decision procedure. While some steps (such as those involved in performing long division) are precisely delineated, many steps involve some form of judgement or interpretation. For example, a decision procedure might involve a duty to aid others when you judge that you can do so at no significant cost to yourself. Alternatively it might only allow killing someone when you judge that they are trying to kill you. Even naïve calculation involves judgements about the goodness of the different possibilities. How are we to interpret perfect execution of a decision procedure such as naïve calculation? Even if the calculating is all done perfectly to the appropriate precision, there are still many ways that one could have judged the goodness of the possibilities. We could assess the procedure on:

(1) the best judgements that could be made,

(2) the worst judgements that could be made,

(3) the judgements that *would* be made, or

(4) 'reasonable' judgments.

Option (1) lends too much flexibility to the decision procedure. For example, a decision procedure that simply tells you to do whatever you judge to be right seems to be intuitively quite poor, yet if the best judgements are made it will lead to quick and optimific action every time. Option (2) would greatly penalise those decision procedures which create many opportunities for intuitive judgments, and this also seems to be incorrect. Option (3) is quite plausible although it takes us very close to the commitment account (discussed in the next section), for it means that only the non-judgment steps of the procedure are being idealised. Option (4) is also plausible: it would need some more fleshing out, but could introduce certain constraints on the judgments, such as freedom from personal bias, freedom from cognitive malfunction and proper use of available information.

Note that this question about how to understand perfect execution of judgments affects the compliance view too. The distinct ways in which we could make a judgment lead to a set of potential actions. We could say that performing the best of these counts as compliance or that performing any of these counts as compliance. Alternatively we could draw upon a notion of idealisation as in (4), saying that performing any of the acts created by 'reasonable' judgments counts as compliance.

The biggest problem with the perfect execution approach is that it does not take the difficulty of executing a certain decision procedure into account. Unless a decision procedure is both very simple and undemanding — such as a handful of simple prohibitions — it is unlikely to be perfectly executed in practice. Sooner or later, the agent is going to make a mistake, succumb to weakness of the will or just decide that the decision procedure is too demanding in the case at hand. The question then is why we would care that a given procedure would be the right one to execute perfectly if it won't actually be executed perfectly.

Suppose that the best decision procedure that Harold can perfectly execute is D_3, but that the good from his actual execution of D_3 will be lower than that which would have come from his actual execution of some other decision procedure, D_4. Suppose further, that we knew that it would actually be better if Harold tried to execute D_4 than if he tried to execute D_3 — perhaps Harold is known to be too weak of will or prone to small mistakes in complex deliberation. In such a case, we should certainly not advise Harold to execute D_3 and we would hope that he in fact chooses to execute D_4 instead. The fact that Harold should perfectly execute D_3 is interesting, but it does not seem to be the most central of our moral considerations.[125]

There is another, similar, problem with the perfect execution account. Suppose that the best decision procedure for Adam to execute perfectly is D_5 and that, unlike Harold, he will actually execute it perfectly. This does not mean that this is the best decision procedure for Adam to accept, for it may be that by trying to execute D_6 (and making occasional mistakes) he could actually lead to better outcomes than he could by perfectly executing D_5. This case seems to be quite realistic, for few people, if any, perfectly execute a decision procedure in the actual world. Perhaps the only decision procedures that we could perfectly execute are

[125] Note that possibilists may still find it to be of relevance, but the objection in the next paragraph holds for both possibilists and actualists.

very simple and very undemanding. It may well be better to attempt to execute more complex procedures even though we somewhat fail in our aims.

For these reasons, it seems that the perfect execution interpretation of the question of decision making is also unsatisfactory.

4.3.3 Commitment

A third approach is to analyse the question of decision making in terms of commitment:

> Q_3: 'Which decision procedure should I be committed to?'[126]

Being committed to a decision procedure involves holding a certain attitude to it. It means that the person is disposed to attempt to make decisions in the way specified by the decision procedure. Due to weakness of will or limitations of cognitive ability, they may not always succeed in using the decision procedure, but they will at least attempt to use it. My use of the term 'commitment' is quite similar to the more common talk of 'acceptance' or 'internalization' of a decision procedure.[127] However there are difficulties with both these terms that I wish to avoid. For example, the verb 'accept' is ambiguous: it can refer to the act of accepting a decision procedure (which has all the problems of the act-consequentialist account of decision procedures), or it can refer to holding an attitude of acceptance to a decision procedure. I take this latter meaning to be equivalent to being committed to a decision procedure but use the term 'commitment' to avoid the ambiguity.

Q_3 suffers from none of the problems that I have so far discussed. Since it does not simply assume compliance, any factors that a decision procedure has which are conducive to such compliance will be properly taken into account in the assessment. Thus decision procedures that actually achieve less good because they are computationally difficult or highly demanding or do not excite moral compunction will be appropriately penalised for this. Because of this, Q_3 picks out D_4 as the decision procedure that Harold should be committed to.

Q_3 also takes into account the other problems that naïve calculation was found to have. A decision procedure that it is good to be committed to will allow the agent to cease deliberating when it would interfere with her concentration or her relaxation. It will allow her to satisfice when the stakes are low and to act quickly in emergencies. Finally, the commitment account takes into consideration the non-act benefits of being committed to a decision procedure. For instance, it might make me happier or calmer to be committed to one decision procedure rather than another and the goodness arising from this is taken into

---

[126] It is worth noting that Q_3 may not be, strictly speaking, a more precise version of Q_1 ('Which decision procedure should I follow?') because being committed to a decision procedure is arguably not a way of *following* a decision procedure. However, the reason I introduced Q_1 was as a formulation of the question of decision making ('How should I decide what to do?') and Q_3 *is* a more precise version of this. Whether or not commitment counts as following is therefore ultimately inessential to the discussion.

[127] See, for example, Brandt (1963), pp. 120–5, and Hooker (2000), pp. 75–80.

account.

Assessing a decision procedure in terms of commitment also allows us to take into account benefits of accepting a decision procedure that can arise merely from the resulting dispositions to act. For example, suppose you are committed to a rule of retaliation to attacks and that you are sufficiently transparent that others are aware of this commitment. In this case, your commitment to the rule may have good consequences even if no-one ever attacks you and you thus never have the opportunity to use this part of the decision procedure. Similar positive effects can come from many other commitments, such as to honesty or faithfulness to a lover.[128]

Finally, interpreting the question of decision making in terms of commitment appropriately takes into account our pre-existing moral beliefs. Sidgwick was correct in stressing the usefulness of our current moral inclinations:

> 'Nor must we neglect the reaction which any breach with customary morality will have on the agent's own mind. For the regulative habits and sentiments which each man has received by inheritance or training constitute an important force impelling his will, in the main, to conduct such as his reason would dictate; a natural auxiliary, as it were, to Reason in its conflict with seductive passions and appetites; and it may be practically dangerous to impair the strength of these auxiliaries.'[129]

Our existing moral beliefs are thus a valuable tool in any attempt to produce good outcomes. They are the product of a considerable amount of moral education by our parents and teachers, and they are backed by strong cultural traditions. Having already committed to them, we can obey their dictates without expending much willpower. Decision procedures that differ considerably from our existing moral frameworks will thus lack this head start and will consequently be subject to more lapses and more effort internalising them. This will have a greater effect on our choice of decision procedures the older we are and the more embedded we are in a context which already has highly structured codes of ethics — medicine or law, for example. This is not to say that we cannot improve upon the status quo — far from it — but the fact that it *is* the status quo is a pertinent fact, and one that Q_3 takes properly into account.

Since it is sensitive to all the considerations that we have searched for in our investigation of decision procedures, Q_3 is the best of these interpretations of the question of decision making. This does not mean that the other interpretations are useless. Indeed, we may often be interested in which decision procedure it would be best to perfectly execute as well as which decision procedure to be committed to, and global consequentialism allows us to ask both questions. However, the commitment account does seem to be somewhat more central to our ethical thought and it is thus the interpretation that I shall focus on hereafter.

---

[128] These points have been discussed in many places: e.g. Lyons (1965), pp. 137 ff.; Williams (1973), pp. 118–122; Pettit and Smith (2000), p. 122; Kavka (1978); Parfit (1984), pp. 20–23.

[129] Sidgwick (1907), p. 482.

4.4 Further clarifications

4.4.1 Which type of rightness

So far in this chapter, I have not explicitly discussed whether the normativity in the question of decision making is best interpreted in objective, belief-relative, or evidence-relative terms. The first thing to note is that within the pluralist position that I support regarding such kinds of rightness, each of these senses provides its own distinct and valid question. At different times, we might be interested in each of these questions.

For example, suppose someone challenges consequentialism on the grounds that if we were to try to follow it, we would *in fact* do worse than we would if we tried to follow some other theory. To this challenge we can reply by pointing out that (global) consequentialism states that the we should (in the objective sense) follow the decision procedure that actually leads to the best outcome. This response shows that the objection was misguided because consequentialism cannot recommend (in the objective sense) a decision procedure that produces a sub-optimal outcome and, moreover, it actively recommends the optimific decision procedure.

In responding to this challenge, it would not be enough to point out that consequentialism states that we should (in the belief-relative sense) follow the decision procedure that leads to the most expected good, because the decision procedure that leads to the most expected good may not lead to the most actual good, and thus this claim of consequentialism does not directly refute the objection. It may well be possible to do so without appealing to the objective version of the question of decision making, but it would require attacking the objection from a different angle — perhaps by arguing that morality is not really concerned with actual goodness, but only with expected goodness. Similarly, if the objection were raised in terms of the belief-relative expected good or the evidence-relative expected good, then these objections can be most directly dealt with by pointing to consequentialism's answer to the corresponding versions of the question of decision making.

However, I do not think that the objective and subjective senses are equally represented in our everyday thoughts about how we should decide what to do. The objectively right way of deciding what to do will typically be quite bizarre. Consider the following example:

> *Symptom X*
>
> Occasionally a patient presents at hospital with symptom *X*. This can be caused by two different underlying conditions: condition *C* and condition *R*. Of these, condition *R* is very rare, accounting for only one case in ten, while the rest suffer from the much more common condition *C*. Both conditions are fatal within one week unless the patient takes the drug appropriate to that condition (conveniently named 'drug *C*' and 'drug *R*'). These drugs are cheap and readily available, but if the patient takes the wrong drug (or both drugs) then there will be no improvement and death will surely result. Unfortunately, there is no way of telling whether the patient has condition *C* or condition *R* and thus no way of telling which drug will save them and which will leave them to die.

This would appear to be a classic example in which the rightness of decision procedures comes to the fore: the right decision procedure for the doctor to accept would be one that guides her to prescribe drug *C* whenever a patient presents with symptom *X*. In the long run, decision procedures that have this feature will produce better outcomes. Since drug *C* is the one that naïve calculation recommends, this consideration could be seen as providing support for the use of a limited form of naïve calculation in the domain of prescribing drugs.

However, consider the following additional information:

> *Symptom X (continued)*
>
> Felicity is a doctor who works at a busy city hospital. During her career, thirty patients present with symptom *X*. True to form, three of them suffer from condition *R*. These people come in on the 18th of July 2020, the 14th of September 2033 and the 9th of December 2050.

Now consider a decision procedure that involves the following condition: if someone presents with symptom *X*, look at the calendar: if it is the 18th of July, the 14th of September or the 9th of December, give them drug *R*, otherwise give them drug *C*. Assuming that no-one else came in on those days of the year, this decision procedure would save all thirty people. So long as Felicity is capable of remembering those three dates, she is also capable of following this decision procedure and thus it (other things being equal) it would form part of the objectively right decision procedure for her to follow.

We could quite rightly maintain that such *ad hoc* modifications cannot be made to all decision procedures, since the decision procedures may become too complex for us to remember them and follow them properly. However, in cases like Felicity's where the stakes are high and the modifications are quite easy to commit to memory, we must admit that it is better for her to be committed to this decision procedure rather than one which always leads to drug *C* being prescribed. Even if she had previously been following the more broad decision procedure: 'use naïve calculation to prescribe drugs to patients', an exception for the three dates would lead to it creating significantly more good.

This pattern of reasoning can be applied outside the medical domain too. Adding a few judicious 'stay home on the 15th of March' clauses to an otherwise normal decision procedure could lead to considerably better outcomes at very little cost in complexity. Consider too the decision procedure: 'buy a lottery ticket with the numbers 18, 7, 79, 9, 12, 80 and donate the winnings to charity'. There is such a decision procedure (with the appropriate numbers) that would allow each of us to do a tremendous amount of good. Indeed, it would probably do more good than most people will actually achieve over their lives.

However, all of these decision procedures are united in being astoundingly *ad hoc*. We could never *know* that it would be best to follow them and they are not what we typically have in mind when we ask 'How should I decide what to do?'. Instead we are more interested in the best decision procedure that we could find by recourse to our existing beliefs or to our available evidence. The sense of 'should' that we are typically interested in is thus a

subjective one. I will not try to adjudicate between the belief- and evidence-relative forms here, as this would depend upon exactly how we flesh out the concept of evidence. For the purposes of this essay, it suffices to point out that while all senses of rightness correspond to valid questions about decision making, the ones that are more central to our understanding of morality are the subjective ones.

4.4.2 Scope

Another point of clarification is the *scope* within which the decision procedure is to be followed. For example, we could can distinguish between the following two questions:

Q_4: 'Which decision procedure should I follow for the rest of my life?'

Q_5: 'Which decision procedure should I follow during the current phase of my life?'

These questions show that the full specification of a role in which to assess a decision procedure requires details about the scope of situations in which the decision procedure is to be followed. In these particular questions, the scope is specified as a period of time. This is not the only way that a scope could be specified. Consider:

Q_6: 'Which decision procedure should I follow in my professional duties?'

This question might be of considerable interest to a doctor who makes use of markedly different decision procedures in the context of her work and her home life.

Arguably, Q_4 is the most prominent of these questions, as it covers the entire remainder of one's life and thus has the largest possible scope. In addition, the decision procedure selected by Q_4 need not be inflexible: for instance, it could have one set of rules for the current stage of your life and another for afterwards, or one set for your professional duties and another for your home life. A small core of rules would then tell you which set to use at which time.

However, as the scope increases, there is an increased chance of it being better to not follow any one decision procedure over that time. This could occur due to considerations of complexity. A flexible decision procedure like the one discussed above is necessarily more complex and harder to remember or follow. And why, in the final part of someone's life, need they still have clauses in their decision procedure to deal with the earlier stages which are now irrelevant? For these reasons, the best decision procedures to follow over one's entire life would probably be self-modifying, allowing the agent to throw away old rules when they are no longer relevant and to create new rules as new information comes to light.

It is clear, however, that Q_5 and Q_6 remain sensible questions and that all three are central to our ethical thinking. Not all scopes are equally relevant — the period consisting of all future Tuesdays would have minimal ethical significance — but there are clearly several interesting scopes that could be considered. Since global consequentialism can allow for them to exist side by side, I shall not attempt to adjudicate between them here, but simply point out that the question of decision making will have different answers depending on the implicit scope.

4.4.3 For one or for all?

In section 3.4.1, I mentioned the possibility of using global consequentialism to go beyond assessment in terms of an individual's choice. By choosing a suitable role, we can come up with the question:

> Q_7: 'Which decision procedure should everyone follow?'

However, this question very quickly runs into a more extreme version of the problem that faced Q_4. It presupposes that there *is* a decision procedure that everyone should follow, but the best achievable outcome may not involve everyone following the same decision procedure. There are certainly advantages of everyone following the same decision procedure, such as an absence of resentment and a greater feeling of camaraderie, but there are also disadvantages such as having to neglect individuals' professions and their different levels of intelligence and strength of will. As in the previous section, one could create a combined decision procedure that begins by asking the agent for their estimates of these values and then proceeding from there, but again there are considerable costs in terms of complexity.

We could, however, proceed with a different question:

> Q_8: 'Which decision procedure is the best for everyone to follow?'

By moving from 'should' to 'best', we no longer need to assume that it would be best if everyone followed the same decision procedure. Even if the optimal outcome involves some people following different decision procedures, there will still be a best decision procedure for everyone to follow. It is just that its being the best may not be all that much of a recommendation.

In the fourth book of *The methods of ethics*, Sidgwick searched for practical rules of conduct for society that could be supported on utilitarian grounds. In doing so, he was attempting to answer a question like Q_8. This focus on communal rules may seem surprising given his individualist account of right action. However, he did allow that it might be better for some people to follow different decision procedures if they were one of 'a class of persons defined by exceptional qualities of intellect, temperament, or character'.[130] It is this consideration which led him to discuss the idea of an *esoteric morality* where certain ethical principles are kept hidden from the public. While he listed many practical disadvantages of an esoteric morality, he considered it a serious empirical possibility that such a system would be best for society and thus that the best way to decide what to do may differ between individuals. However, he seemed to consider this as a matter of minor interest and focused on finding the decision procedure whose universal adoption would be optimific.

There are two key reasons why we might be interested in individual choice of decision procedures. The first is present in the idea of an esoteric morality: the best overall pattern of

[130] Sidgwick (1907), p. 489.

decision procedure following may involve different people following different decision procedures. The second concerns what we are to do when some people don't follow their optimific decision procedures: should we follow the decision procedure that was optimific assuming everyone played their part, or should we follow the decision procedure that will be optimific relative to the world as it is? The difference between these two options was not very large in the examples that predominantly occupied Sidgwick. To a near approximation, you should follow a decision procedure involving honesty, promise-keeping, compassion and so forth regardless of whether others do. However, not all cases are like this.

Consider the great good that can be done by donations to effective charities working in very poor countries. If everyone in the richer countries were to do their part, then we would not need to give very much each. The best decision procedure for everyone to follow would thus make quite modest demands of each of us, say 2% of one's annual income for each person in the richer countries. However, the overwhelming majority of people are not doing their part, and thus a great good could be achieved by an individual if she were to give significantly more than her fair share, such as 20% or 50% of her income.[131] The best decision procedure for an individual to follow is thus dramatically changed by the non-compliance of others.

As this shows, there is a great difference between Q_8 and the individualistic rendering of the question of decision making. Without a doubt, both versions are natural and important moral questions, but they do have quite different focuses. Since the focus within act-consequentialism is on what is right for an individual given the world as it actually is, and since it is the generalization of this question that has lead us to consider the question of decision making, I shall continue to interpret it in an individualistic sense in this essay.

Before we move on to the next chapter though, I should note that there are also many interesting related questions which feature a restricted universalisation. For example:

> Q_9: 'Which decision procedure is the best for doctors to follow during their work?'

This is asking for the best code of conduct for doctors and is a standard question of practical ethics. It combines a restricted universalisation with a restricted scope of application. It is too specific to be what is commonly asked when people wonder how they should decide what to do, but it shows the versatility of global consequentialism and the decision procedure approach that it can also pick out and answer this question.

Chapter summary

In this chapter, I investigated numerous ways in which we could interpret the question of decision making. I have shown that decision procedures are best interpreted as consciously

[131] This matter is very large in absolute terms. For example, if someone earning £30,000 per annum were to give 50% of this to the Fred Hollows Foundation, they could cause around 20,000 people to be cured of blindness over their life (a conservative estimate based on extensive analysis of the foundation's annual reports). This is likely to be a much larger benefit than the individual would produce in the other parts of his or her life.

applied mental processes for forming intentions, that they can involve intuitive steps, that they can make use of repugnance and compunction, and that they can allow room for the agent to choose freely in some domains of choice. I have also looked at the question of granularity and have shown that it is unclear which level of granularity is most relevant.

I have shown that decision procedures are most usefully assessed in terms of commitment, rather than compliance or perfect execution, that the question of decision making is most usefully asked in terms of subjective rightness, and that decision procedures can quite naturally be assessed over a variety of different scopes. While one could usefully ask which would be the best decision procedure for everyone (or nearly everyone) to follow, I have shown that the question of which decision procedure for an individual to follow can have a very different answer and is the more natural question to ask for those concerned with the fundamental consequentialist intuition.

In the next chapter I perform a similar analysis for the question of motivation.

Chapter 5

Motivation and Character

In the last three chapters I have explained how consequentialism can be applied to decision making. In Chapter 2, I showed that the traditional forms of consequentialism could not adequately assess decision procedures. In Chapter 3, I showed how a recent formulation of consequentialism — global consequentialism — could directly assess acts and decision procedures (and everything else). In Chapter 4, I explored the nature of decision procedures and the ways in which we can assess them.

Now that I have analysed decision making in some depth, I shall turn to the other focus of this essay: motivation. Many of the same questions that were asked about decision making can be asked about motivation, but this time we can proceed much faster. Section 5.1 is an echo of Chapter 2, returning to the question of whether the traditional forms of consequentialism suffice, but this time looking at whether act-consequentialism and motive-consequentialism can meet our needs. I conclude that they are at best partial successes, and that we should once again turn to global consequentialism.

Since Chapter 3 applied to motivation just as much as it applied to decision making, it requires no analogue here. The remaining sections therefore revisit the questions of Chapter 4, exploring how global consequentialism can best be applied to motivation. In section 5.2 I briefly examine the nature of motivation and its relationship to character. I conclude that both motivation and character are central focal points, and that they can be usefully assessed as individual motives or traits, and as entire patterns of motivation or characters. In section 5.3, I explore several questions about the most central roles in which to assess motivation or character, and in section 5.4 I clarify some miscellaneous aspects of this assessment.

5.1 Do the traditional forms of consequentialism suffice?

5.1.1 Act-consequentialism

The possession of a certain set of motives is not in itself an act. Therefore act-consequentialism does not directly assess the possession of motives. However, there are three senses in which act-consequentialism could be said to indirectly assess motives.

The first of these involves looking at the motives behind our acts. If a doctor delivers a lethal injection while motivated by compassion, we may want to say that this is a different act than were she to deliver the injection out of hatred for the patient. Indeed we may want to say this even if it involves an identical bodily action at precisely the same time. Just as one can specify an act in a manner that ties it to a particular time, place and agent, so one could say that a complete specification of the act includes the motivation of the agent. If so, this provides a sense in which act-consequentialism can assess motives. For example, suppose the act with the best consequences in a given situation is one in which a man helps his wife out of love for her. In this case we could say that the right act is that of helping out of a motive of love, and this provides some kind of commendation for the motive of love.

While this strategy does provide a form of assessment for motives, it too weak for our purposes. Intuitively, we would like to be able to assess motives such as love in more general terms. It is not enough to know that the right act involves love in situation *A* and hate in situation *B*. We want to be able to determine which of these motives is better in a more general sense; to find out what the best motives would be in the longer term, not just in one particular situation. The assessment of motives as parts of acts thus suffers from the same problem that we saw in section 2.1.1 with the assessment of the act of making a decision.

A second option for act-consequentialists is to directly incorporate a value for motives into the axiology. Thus, it might be non-instrumentally bad to hate and non-instrumentally good to love. This is a very different strategy as it is not assessing the motives in terms of their consequences, but in terms of their own value. By changing the axiology in this way, the theory would be very different from theories like utilitarianism in which only a single type of thing (such as happiness, or desire satisfaction) has non-instrumental value, and where the value of other things depends upon whether they lead to this central source of value. However, consequentialism is a broad church and is open to theories with complex axiologies such as this.

There are three problems with this approach to assessing motives. One of these is that it raises questions about why particular motives are good and others are bad. It is possible to have a consequentialist theory which offers no explanation for its judgments about which motives are good and bad, but this would be at the very least unsatisfying and at worst *ad hoc*. It seems preferable if the theory can provide a unifying account of why certain motives are good and bad. One of the most obvious explanations — that motives are good or bad insofar as they lead to good or bad outcomes — is not available. There is thus an explanatory burden that adherents to this approach would need to discharge.

Another problem concerns cases where having a certain good motive would lead to a very bad outcome. Suppose that possessing the motive of benevolence is considered good by the axiology, but in a certain circumstance it would knowably lead to disaster (and thus the badness of its consequences would outweigh the goodness of its possession). If we do not provide a direct consequentialist assessment of motives, then we cannot take this disastrous result into account, even though it seems to be very relevant to a proper assessment of the value of benevolence in the case at hand.

The final problem is that this approach is not very plausible for assessing decision procedures. There is some intuitive appeal to saying that love or compassion is non-instrumentally valuable, but little intuitive appeal in saying that the following of certain decision procedures is non-instrumentally valuable. Thus, if we want to assess both motives and decision procedures, this approach does not seem adequate. In contrast, the global consequentialist approach works cleanly for every focal point, does not require modifying the axiology, and remains true to the fundamental consequentialist intuition.

The third option for act-consequentialists is to assess the act of inculcating a motive. This option is very similar to that of assessing the act of inculcating a decision procedure (see section 2.1.4). All of the same comments are applicable here. This approach will indeed indirectly assess our motives and will assess them 'in the long run' since the gain (or loss) of a motive will affect many future acts. Unfortunately, there is a degree of disconnection between the act of inculcation and the effects of the motive, just as there was for inculcation and decision procedures. For example, it might be best for me to start inculcating a motive now, but only because it will take a long time before I gain the motive. There are also problems when there is no act that is both necessary and sufficient to inculcate a motive, for then there is no simple connection between the rightness of a particular act and the commendation of the motive.

I think that this third option is ultimately the best way of assessing motives within act-consequentialism. However, the indirection involved means that it is only a partial success. Since global consequentialism provides a more direct form of assessment, does not leave a gap between the act and the motive, and is arguably a simpler theory overall, one would require a good reason to reject it in favour of (mere) act-consequentialism. As I shall show in Chapter 6, such a reason has not yet been provided.

5.1.2 Motive-consequentialism

Let us therefore turn our attention to motive-consequentialism. Outlined by Robert Adams in 1976,[132] motive-consequentialism would seem to be an excellent candidate for a non-global form of consequentialism that can fully and directly assess motives. Adams introduces motive-consequentialism with the criterion:

> '…one pattern of motivation is morally better than another to the extent that the former has more utility than the latter.'[133]

This suggests that motive-consequentialism is only concerned with evaluative assessment of patterns of motivation, and not with normative assessment. However, there are also a number of remarks that Adams makes later in the paper in which he applies normative terms to patterns of motivation. For example, he says of a person that 'his motivation is right by

[132] Adams (1976).

[133] Adams (1976), p. 470.

motive utilitarian standards',[134] and later he remarks that motive-consequentialism is 'about what patterns of motivation one ought to have.'[135] It thus seems that Adams also takes motive-consequentialism to include a criterion of rightness for motives, where the right motives are those that maximize goodness.[136]

Motive-consequentialism is thus a theory about the betterness (and possibly the rightness) of patterns of motivation in terms of their consequences. This makes motive-consequentialism a sub-theory of global consequentialism. In other words, every claim made by motive-consequentialism is also made by global consequentialism (though not vice versa). In particular, both theories make the same claims about patterns of motivation, and thus motive-consequentialism is exactly as good as global consequentialism at assessing them.

Let us, however, turn to motive-consequentialism's assessment of acts. Unlike rule-consequentialism, motive-consequentialism does not include a principle defining right action in terms of best motivation. Adams considers such a principle, and also considers using the act-consequentialist criterion of rightness to assess acts directly in terms of their consequences. Ultimately, however, he finds both options unsatisfactory and doesn't specify any particular way for motive-consequentialism to assess acts.[137] Instead it is just a theory about motivation and remains silent on questions of acts. To judge acts as well, one would have to hold motive-consequentialism in conjunction with a separate theory pertaining to acts.

Thus, on its own motive-consequentialism will not suffice for the purposes of this essay. It answers the third of our key questions — the question of motivation — but remains silent on the question of action and the question of decision making. We could, of course, supplement motive-consequentialism with theories to deal with these areas, but this will not really help. If we do not directly assess these focal points in terms of their outcomes, then we violate the fundamental consequentialist intuition, and I have been approaching this entire essay from the perspective of a hypothetical act-consequentialist who is motivated by this intuition and would like to assess decision making and motivation without violating it. Alternatively, if we add principles to directly assess acts and decision procedures in terms of their consequences, then we have a theory that seems strictly worse than global consequentialism, for it is more contrived, more complex, and less expressive, while subject to all the criticisms that have been levelled against global consequentialism.

5.2 Motives and character

Many consequentialists have been interested in applying the consequentialist test to motivation (broadly construed). This is a very important focal point for two reasons. Firstly, it

[134] Adams (1976), p. 471.

[135] Adams (1976), p. 474.

[136] See also Feldman (1993), p. 203 on this point.

[137] In section 6.1.1 I shall explain Adams' main objection and why I do not find it damaging to global consequentialism.

is pivotal in guiding us to good action and is more or less within our control. Anyone interested in producing good outcomes would therefore do well to pay considerable attention to motivation. Secondly, it is a standard subject of moral assessment in common-sense morality and other normative theories such as Kantian ethics and virtue ethics.

However, consequentialists who are interested in these matters are often unclear about exactly what it is that they wish to assess. So far, I have spoken of 'motivation' and 'motives' and have been using these terms in a very broad way. However consequentialist also write extensively of 'character' and 'character traits'. Some authors carefully distinguish between motives and character traits, focusing their attention on one or the other. For example, Adams writes:

> 'By "motives" here I mean principally wants and desires, considered as giving rise, or tending to give rise to actions. A desire, if strong, stable, and for a fairly general object (e.g., the desire to get as much money as possible), may perhaps constitute a trait of character; but motives are not in general the same, and may not be as persistent, as traits of character.'[138]

Other authors use one or both of these terms without any precise explanation of whether they are taken to mean different things, or what (if anything) the difference is taken to be. For example, Mill writes about consequentialist assessment of character in his *System of Logic*, and about motives in *Utilitarianism*, but it is not clear how we are to understand the relation between them.[139]

In this essay, I shall not attempt to determine exactly how we should use these terms. For one thing, it does not matter very much to the present project: we are concerned with assessing different focal points, not with naming them. However, it will be useful to note some differences in how these terms are commonly used.

Motives and character traits are both aspects of a person's psychology. They are psychological features which are distinct enough, and common enough, to warrant individual names such as compassion or greed. At a minimum, motives and character traits involve certain psychological dispositions (or sets of dispositions). However it is possible to conceive of them in much richer ways, and they may also be said to involve many other psychological features such as emotions, desires, perceptions, attitudes or interests. The term 'motive' literally means something that moves us to action.[140] There thus seems to be a particularly close link between motives and desires, though I shall not attempt to determine the exact nature of this link.

---

[138] Adams (1976), p. 467. This distinction is also used by Driver (2001), pp. 89–95, who chooses instead to focus on character traits.

[139] Mill (1843), Bk. VI, Ch. 12, §7, and Mill (1861), Ch. 2.

[140] Although as Bentham has pointed out [(1789) Ch. X, §I, par. III.], motives can also serve to restrain us from action.

As Adams remarks, the term 'character trait' is typically used when the psychological feature is particularly stable and long lasting. A man might have greed as his motive at one moment, but not at the next. However his character traits are not so fickle: the character trait of greediness is usually gained or lost quite gradually. Perhaps, as Julia Driver suggests, the connection between the two is that character traits involve the disposition to be motivated in a certain way at certain times.[141] Thus, a greedy man (one with the trait of greediness) is disposed to be motivated by greed in many situations, even though he need not always be motivated by it.

Another distinction in our common usage of the terms is that motives can be more specific than character traits. Someone's motive in a particular instance might be to avoid being caught before they could dispose of the evidence, but this seems far too specific to be a character trait. We won't typically call something a character trait unless it is quite general in its applicability.

There is undoubtedly a connection between motives, character traits, and *virtues*. For example, the term 'compassion' can be used to designate a motive, a character trait, and a virtue. Driver identifies the virtues as a subset of character traits: those that normally lead to good outcomes.[142] Whether or not one agrees with her particular consequentialist identification of virtues, it does appear that in ordinary usage, we use the term 'virtue' to cover certain clusters of dispositions that are fairly long-lasting and characteristic of a person. It thus appears to align more closely with our use of the term 'character trait' rather than our use of 'motive'.

There is also an important distinction which is orthogonal to these. Sometimes it is convenient to talk about an individual motive, and at other times to talk about a person's pattern of motivation. Similarly, we can talk of an individual character trait or of a person's character. This is a distinction between individual entities and their aggregates. As we shall soon see, there is a question as to whether our consequentialist assessment should be at the individual or the aggregate level.

We could thus represent these rough distinctions as follows:

	individual	**aggregate**
transitory	motive	pattern of motivation, motive-set
lasting, general	character trait	character
lasting, general, positive	virtue	

Table 2.

We now come to questions about whether some of these categories are more central than others in moral life. Let us first look at the distinction between the individual and the

[141] Driver (2001), p. 92.

[142] See Driver (2001), p. 82, for her more precise definition.

aggregate. We know how we can assess patterns of motivation and characters. Actualist versions of global consequentialism look at what would happen if a given person had that pattern of motivation or that character. Possibilist versions look at which pattern of motivation or character the person would have in the best world open to them. However, there are a number of problems which arise when we try to assess individual motives or character traits.

In order for actualist global consequentialism to compare two things in a particular role, they need to be alternatives to each other in some sense. At the very least, they need to be what I have called 'soft alternatives': if either thing were to be in the given role, the other would not be. This will not always be the case for motives or character traits. For example, suppose we were to compare the character traits of compassion and generosity in the role of being possessed by a particular individual (Matthew) over a particular period. It may very well be that Matthew is already compassionate and will be compassionate over that period whether or not he also becomes generous. In such a case, we cannot perform an actualist comparison. We will still be able to perform actualist comparisons in certain cases (for example it is typical that if someone is compassionate they are not also malevolent) but we won't always be able to do so. This inability to perform some comparisons makes it incredibly difficult to determine which is the best or the right character trait (or motive), for this requires that the trait (or motive) is compared to all others, and thus that it is at least a soft alternative to all others. It will thus be very rare that such a comparison can be made.

This is not necessarily a bad thing, for there is something quite strange about asking whether, say, compassion is better than generosity — particularly in situations where they are not real alternatives to each other, but could both be possessed. The same holds for asking whether compassion is the best character trait (or motive) in like situations. Arguably, actualist global consequentialism's refusal to answer these questions is to be commended, for the questions border on incoherence.

There is, however, another type of comparison that actualist global consequentialists *can* make, and it is one that seems more natural for the situation. They can ask whether it would be better to possess the character trait of compassion or to not possess it.[143] These options are true alternatives (in both of my senses) and we can determine the answer by comparing the outcome that would arise were someone to possess the trait to the outcome that would arise were they not to. Indeed, if we use the actualist proposition-based global consequentialism of section 3.3.2, then we can even say that Matthew ought to possess the character trait of compassion over not possessing it (so long as this is within his power and would produce a better outcome). We could even say that he ought (*simpliciter*) to possess the character trait of compassion (so long as the above circumstances are met and the relevant alternative set is taken to include only one option: that Matthew doesn't possess the character trait).

The situation is similar, but slightly different for possibilist theories of global consequentialism. These do not assess whether one trait or motive is better than another, or whether a trait or motive is the 'best', but they do consider the relative merits of possessing

[143] They can also ask *how much better* it would be.

the trait or motive versus not possessing it. These theories say that someone ought to possess a certain character trait (or motive) if it is within her power and she possesses it in the best world open to her.

Thus, both actualist and possibilist forms of global consequentialism can provide certain kinds of assessment of individual traits and motives. However, for the actualist theories this assessment is quite different to the assessment of acts or decision procedures or characters or patterns of motivation, since the objects of evaluation here are not all alternatives to each other and thus there is no 'right' character trait or motive, only traits or motives that it is better to possess than not to possess, or that one ought to possess.

Having thus outlined the complexities of assessing individual traits or motives, let us return to the question of whether their assessment is more or less central than that of their aggregates: patterns of motivation and characters. Firstly, it is clear that both have been found worthy of attention. Adams and Parfit explicitly focus on patterns of motivation, while Bentham focuses on the merits of individual motives. There is an entire ethical tradition around studying the individual virtues (or putative virtues), and both Sidgwick and Driver have spent considerable time assessing individual virtues in terms of their consequences.

An obvious problem with a consequentialist assessment of individual traits or motives is that their effects are not independent of each other. For example, it may be very good to have a certain trait when it is part of one overall character and bad to have it when it is part of another. However in practice this effect is not that large, and both Sidgwick and Driver have been able to make sensible judgments about the consequences of a number of character traits considered in isolation. The issue of their interdependence is more a theoretical concern, complicating the question of just what it could mean to have a precise assessment of each trait or motive, but having relatively little effect on the big picture.

In the end, it appears that questions on both the individual and the aggregate level are central: questions about individual traits or motives and questions about patterns of motivation or characters. Both types can be asked, though there are some complexities with the assessment of individual traits and motives, both types are clearly of interest to someone trying to better promote the good, and both types have received particular attention from consequentialists and nonconsequentialists.

Let us now turn to the question of whether motivation or character is more central to ethical thought. Again, there are consequentialists who have focused on each of these and both are clearly of significant importance for someone interested in promoting the good. If forced to choose between them, then the fact that character traits are more stable and lasting than motives would seem to give a slight upper hand to questions of character. However, we are not forced to choose. Both are clearly of considerable importance and global consequentialism can be used to assess both.

This is a considerable advantage of global consequentialism over consequentialist theories that have only one focal point, such as motive-consequentialism. Global consequentialism allows us to assess both motives and characters, both individually and in aggregate.

Moreover, it means that we do not need to completely untangle the questions of just what a motive is and just what distinguishes it from a character trait. Where single focal-point theories require a canonical description of their focal point, we can use global consequentialism to assess motives and character traits under any particular definitions that one cares to specify. Obviously there is good reason to think carefully about these definitions and choose ones that are most relevant to ethical thought, but we need not go any further down this road in the present essay.

There is a contrast here between the question of decision making and the question of motivation. In Chapter 4, we found that we could make considerable progress towards clarifying the question of decision making: there was, more or less, a single most central conception of decision procedures and a single most central role in which to assess them. There were some complications, but it still felt as if there was a single question that was being clarified. However, in examining the nature of motivation, we have already found four separate interpretations that are central to ethical thought. The 'question of motivation' is thus not so much being clarified as fragmented into a number of important but distinct questions on intimately related topics. This is not a problem in a global consequentialist approach, but is certainly worthy of note. I will sometimes continue to talk of 'the question of motivation', but from now on, this is to be understood as shorthand for a small number of related questions about motivation and character. In the same vein, I shall sometimes illustrate my claims with reference to just one of these focal points when I judge that it would be overly cumbersome to refer explicitly to all four.

5.3 Roles

Global consequentialism can only assess a focal point with respect to a given role. In which roles then are we to assess motives, character traits, patterns of motivation and characters? The choice of role is much simpler for these focal points than it was for decision procedures. In that case there were several prominent contenders: compliance, perfect execution, and commitment. Of these, commitment was found to be the most central to our ethical thought.

In the present case, the central role seems more obvious: possession. Motives, character traits, patterns of motivation and characters are all things that someone can *possess* (or *have*). To possess a character trait means that your psychological features include those involved in that trait. To possess a motive means something very similar, though possibly with different types of psychological features. Possessing a character is like possessing a character trait, but in addition to requiring that your psychological features include those of the specified character, it also requires that you have no further character-traits. In other words, it requires that your character is an instance of the specified character. Possessing a pattern of motivation is like possessing a motive, but again, it requires that you have no additional motives beyond those specified.

Possession is a very natural role in which to assess all of these focal points, and it is difficult to find any serious rivals. I can only think of one and it is rather difficult to develop precisely. I shall, however, provide a rough account of it and this should be enough to show that even if

it can be developed more precisely, it is a less central role than possession.

Instead of looking at what would happen if someone possessed a given character, we could look at what would happen if they acted in strict accordance with that character. This role is intended to be analogous to the compliance account of decision procedures. Both abstract away the details of the agent's actual psychological state (whether they are committed to the decision procedure, or possess the character) in favour of whether their actions are in strict accord with the decision procedure or character.

For example, suppose that we want to assess a particular character that involves truthfulness as one of its chief traits. It might be possible for someone to possess this character without always acting in strict accordance with it. Perhaps they sometimes suffer from a weakness of the will. We would still say that someone possessed this character even if there were a handful of occasions on which they told unimportant lies or made minor misleading statements. However, when the person does these things, they would not be acting in strict accordance with the character.

Some people may not accept this distinction. They may say that someone can't help but act in strict accordance with their character. If so, then the roles of possessing a character and acting in strict accordance with a character are one and the same, and we can harmlessly use the former. I think that it is quite plausible, however, that one can act 'out of character' and thus that the outcome of possessing a character could be different to that of acting in strict accordance with it.

Let us therefore suppose for the sake of argument that one can act out of character. In this case, which of the roles is more central? One difference between them is that it is, in a sense, more demanding to act in strict accordance with a character than to merely possess it. This can come into play when we consider the best character to possess or to act in strict accordance with. For example, suppose that best character for Charlotte to possess is one that involves a high degree of benevolence. This causes her to donate a lot of her salary and produce a very large amount of good. She does not always act benevolently — she has her selfish moments — but on the whole she is much more benevolent than average.

Let us now consider the best character for Charlotte to act in strict accordance with. It may not be psychologically possible for her to always act with a high degree of benevolence. The best character to act in strict accordance with would therefore be one that did not demand this.[144] Instead it would either have a smaller role for benevolence that she could always live up to, or no place for benevolence at all. However, in either case we can suppose that Charlotte would not be able to produce as much good. By restricting our range of characters to those that Charlotte can act in strict accordance with, we have removed from consideration the one whose possession would lead to the best outcome. When there is such a

[144] At least this is true when we restrict the set of possibilities to those that are possible for the agent in the sense of 'ought' implies 'can'. I believe that this is the natural way to interpret intuitive judgments about the best motives for one to have. This can be made technically precise by modifying roles in the form *is a character that Charlotte acts in accordance with* to *is a psychologically possible character that Charlotte acts in accordance with*.

conflict, the more relevant question appears to be about which character it would be best for her to possess: why focus on setting a standard low enough that it can always be met, when a higher standard can produce a better outcome?

A second difference between the two roles concerns the effects that a character can have other than through our acts.[145] It may be that possessing a certain character makes one calmer or more happy. By focusing on whether a person acts in strict accordance with a character, these other benefits are not assessed. Possessing a character could also have effects on other people without being mediated by acts. For example, it might cause you to blush in a certain situation, or to unconsciously adopt a different facial expression in reaction to unexpected news. The ability to take into account these features of a character is another advantage of focusing on the consequences of possession of a character rather than those of acting in strict accordance with it. I therefore conclude that possession is the more central role for assessing character, and similar considerations show that it is also the more central role for the assessment of character traits, motives and patterns of motivation.

The choice between possession and strict accordance does not completely determine a role. For example, we may have decided to assess the possession of a particular pattern of motivation by a particular person, but we have not yet specified the scope over which this motivation is to be possessed. We may be interested in the effects of possessing a pattern of motivation over a very short period, such as during the course of a certain conversation, or over a long period such as the remainder of one's life or a stage of one's life. We may also be interested in assessing the possession of a pattern of motivation during one's professional work, or when in the presence of a certain person. These latter possibilities constitute a series of temporally disconnected periods, but form coherent scopes over which we could usefully assess patterns of motivation. The choice of scope also affects the assessment of characters and of individual traits or motives. However, since characters and character traits are considered to be relatively long lasting and slow to change, the shorter scopes have less relevance to their assessment.

Only the largest scope — the rest of one's life — could have any claim to be the canonical scope for the assessment of motivation. However, there are clearly times when each of these scopes are particularly salient, and it seems that a great many of our intuitive assessments of motivation fall under one of the smaller scopes. Since global consequentialism can ask about any of these scopes, I shall again refrain from declaring any of them to be central, but merely stress that the scope is an important part of the role in which we assess motivations or characters and thus that careful assessment of these focal points will involve consideration of the relevant scope.

Finally, we could assess characters (or patterns of motivation etc.) in the role of being possessed *by a given individual*, or in the role of being possessed *by everyone* (or almost everyone, or the members of a given society, or medical practitioners…). Unlike in the assessment of rules or decision procedures, there is not much emphasis in the literature on assessing characters in a collective way. There may sometimes be an assumption that the best

[145] See, for example, Williams (1973), p. 118–135.

characters identified in each of these ways will converge, but there is little explicit argument that both: *(i)* the two approaches diverge and *(ii)* we should use the collective approach.

At one point, Crisp does suggest something like this.[146] He was considering McCloskey's famous case of a sheriff who could prevent a destructive riot only by framing an innocent person.[147] Crisp noted that in this case it would be bad for the sheriff to possess the character trait of justice, for that would prevent him from averting the disaster.[148] However, Crisp notes that this is an aberrant case and when assessing the merits of the trait of justice, we should consider what would happen if no sheriffs possessed it. This appears to support a collective assessment of character traits, but I do not think that this could be his true view, for on the very next page he remarks upon the merits of taking account of each individual's intellectual, motivational and reflective capacities when determining which traits will be best for them to possess. It thus appears that he was merely using his example where no sheriffs possessed the trait of justice as a thought experiment to show that we have very little reason to believe that a lack of justice would be beneficial in any particular situation. In other words, he was using the example to make the distinction between whether justice will have objectively good effects or whether it will have good effects in the evidence-relative or belief-relative senses.

It may well be very useful in some situations to ask about the best characters for some set of people to have. However, we must bear in mind that such assessments make certain abstractions. As Crisp noted, they are insensitive to the individual's attributes. The best causally possible set of characters for the members of a group might involve different characters for different people. In this case, why bother asking which single character would be best for all to have? Secondly, the best character for an individual to hold depends a great deal upon whether we are idealising the characters of others. For example, in reality it is best that I am not completely trusting, but the best character for everyone to have may well involve being completely trusting. Such issues are familiar from the study of rule-consequentialism, but are worth mentioning here to remind us that they would also arise when we ask questions about the best characters (or patterns of motivation etc.) for a given group of people to possess. Global consequentialism can indeed ask the collective versions of this question, but it appears that to someone who shares the fundamental consequentialist intuition, they will be of less interest than the individualistic version.

---

[146] Crisp (1992), p. 157–158.

[147] McCloskey (1963).

[148] Crisp actually phrases his considerations in terms of possessing the virtue of justice, which I think is best understood as an assessment of a particular character trait. Otherwise one is forced to use apparently paradoxical language, when considering that a particular virtue may not really be a virtue. It seems much better to render this thought as the possibility that a character trait which is thought to be a virtue may not really be a virtue.

5.4 Further clarifications

There are three further issues that I wish to briefly explore. The first of these is whether we are to assess patterns of motivations (and character etc.) objectively or subjectively. The relevant considerations are very similar to those involved in assessing decision procedures,[149] so I shall be brief. As with decision procedures, we may sometimes be particularly interested in the objective assessment of motivation. For example, if someone introduces the 'paradox' of benevolence by saying that being motivated purely by universal benevolence would actually lead to a bad outcome on consequentialism's own terms, then we can most directly refute this criticism by pointing out that consequentialism tells us that we objectively ought to possess whichever pattern of motivation it is that leads to the best outcome.

However, we are typically more interested in the motivations (or characters) that lead to the best outcome in one of the subjective senses, as these reflect the limited epistemic situation from which we make all of our choices. These senses also have the property that there are fewer times in which intuitively bad motivations are judged right or best. For example, it was subjectively right for McCloskey's sheriff to possess a strong motivation to uphold justice. In other words, at any stage sufficiently far in advance of the fateful situation to allow the sheriff to deliberately change his motivations, the balance of evidence and belief would have suggested that it was very good for him to keep his devotion to justice. I therefore hold for motivations and character what I held for decision procedures: the subjective senses are the most central, but the objective sense is occasionally useful too.

The remaining two issues arise in relation to Driver's consequentialist account of virtue.[150] After considerable discussion, she decides upon the following connection between virtues and consequences:[151]

> x is a virtue *iff* it is a character trait that produces actual good consequences overall or systematically.

From the perspective of global consequentialism, there are two strange features of this account. The first is that, as Driver makes clear elsewhere in the book, it demands only that a character trait produces a net balance of good. That is, the outcome would be better were the agent to possess it than were she not to. This is a *satisficing* account of virtue. In my account of global consequentialism, I specify rightness and ought in terms of maximization and provide no special role for satisficing. However, as mentioned in the analysis of character traits in section 5.2, one *can* use global consequentialism to ask whether it would be better were something the case than were it not. Such questions have no privileged status in the theory, but this does not mean that they are uninteresting or irrelevant.

The surprising thing about Driver's account of virtue from the perspective of global

[149] See section 4.4.1.

[150] Driver (2001).

[151] Driver (2001), p. 95.

consequentialism is thus not that it has a place for satisficing, but that it stops there and doesn't go on to grant a privileged role to the best combination of character traits. However, on closer examination this puzzle is resolved. While her criterion of what is and is not a virtue relies upon satisficing, Driver adds that:

> '…the better virtues will be the ones that produce more good. … If there is a virtue that produces more good than any other, then that would be the best.'[152]

We can thus see that her formula is satisficing because it is just a criterion for whether or not a given character trait is a *virtue*, and is not intended to encompass all her views on the consequentialist assessment of character traits. This makes sense, as no doubt the only way for a consequentialist account of virtue to bear much relation to pre-theoretic use of the term is through satisficing. We must simply take care not to confuse this with the view that consequentialism is content with suboptimal combinations of character traits or motives.

The other strange feature of Driver's formula is that it isn't concerned with the consequences of a particular person possessing the given trait, or the consequences of everyone possessing the trait, but with the *systematic* production of good consequences.[153] This means that a virtue need not produce good outcomes for everyone (perhaps some people never end up in a situation where the virtue comes into play) and the virtue can even produce bad outcomes for some people (perhaps through bad luck or because others take advantage of their virtue), but there must be a systematic connection between the possession of the virtue and the production of the good.

This use of 'systematic' would fit into my account as an alternative sense of rightness or goodness, alongside the objective sense, the belief-relative sense, and the evidence-relative sense. Interestingly it could not be classified as subjective as it is not relativised to a given agent, or set of evidence, or something similar. Instead, it would be another objective sense. While this does hint at some interesting possibilities, I shall set it aside as it too seems to be brought in for the purpose of providing a consequentialist criterion of virtue which fits the intuitive conception. When it comes to consideration of the best motivation or character for someone to possess, I continue to hold that the standard senses are the most central.

Chapter summary

In this chapter, I investigated the consequentialist assessment of motivation and character. I first examined the different ways in which they could be assessed within act-consequentialism and motive-consequentialism, finding that none were entirely satisfactory — especially when we also want to be able to assess decision making and action. I then looked briefly at the nature of motivation and its relationship to character, showing that there are central consequentialist questions concerning each of these, as well as the assessment of individual motives and character traits. Following this I examined a number of questions concerning

[152] Driver (2001), p. 74.

[153] Adams also considers the systematic production of good outcomes [Adams (1976), pp. 479–480].

the appropriate role in which to assess these focal points. I found that it is more natural to assess the possession of a pattern of motivation or character than to assess acting in strict accordance with it, that there are many important scopes over which we might assess a pattern of motivation, and that we are primarily interested in questions concerning the motives or character of individuals rather than groups. Finally, I examined some miscellaneous issues, showing that once again we are primarily interested in the subjective senses of goodness and rightness, and that we do not need to worry very much about satisficing or 'systematic' accounts of consequentialist motivation or character.

Chapter 6

Problems addressed

In this chapter, I shall address four important objections to applying consequentialism to decision making and motivation. Namely:

(1) the inconsistency objection

(2) a regress concerning decision procedures

(3) the indistinctiveness objection

(4) the inappropriate assessment of blame and guilt

Of course the form of consequentialism that I advance is still open to many of the classic objections to act-consequentialism. These have been discussed at great length in the literature and continue to provoke debate. Unfortunately, I cannot devote the space to do them justice in this essay, and they would in any event distract from the task at hand, which is to show how consequentialism can be successfully applied to decision making and motivation. I therefore restrict my attention to objections that are distinctively aimed at this project.

6.1 The inconsistency objection

The most prominent objection to applying consequentialism to everything is that judgments about one focal point can sometimes come into conflict with judgments about another. I call this the *inconsistency objection* and in this section I shall explore its various forms, showing that it is not a serious obstacle to the global consequentialist project.[154]

[154] There is a large literature related to the inconsistency objection and I shall not be able to discuss it all in this section. Much of it is phrased in terms of 'blameless wrongdoing' and is focused on finding ways for consequentialism to avoid such conflicts, rather than on showing why they are a problem in the first place. I shall therefore focus on those parts of the literature that are most relevant to the present project. Good sources that focus on avoiding the need for 'blameless wrongdoing' include Gruzalski (1986), Mason (2002), Braddon-Mitchell (2004) and Louise (2006).

6.1.1 Adams and Hooker

Robert Adams describes the root of the problem as follows:

> 'It can be better, by motive-utilitarian standards, to have a pattern of motivation that will lead one to act wrongly, by act-utilitarian standards, than to have a motivation that would lead to right action.'[155]

In light of this, he ultimately concludes:

> '…it is doubtful that direct application of the test of utility to everything results in a system that counts as a morality.'[156]

Brad Hooker presents the objection along similar lines, though he focuses on the conflict between decision procedures and acts:

> 'Suppose, on the whole and in the long run, the best decision procedure for you to accept is one that leads you to do act *x* now. But suppose also that in fact the act with the best consequences in this situation is not *x* but *y*. So global consequentialism tells you to use the best possible decision procedure but also not to do the act picked out by this decision procedure. That seems paradoxical.'[157]

Both of these explanations cast the problem in terms of something that is morally recommended leading to something that is wrong. For Adams it is the better motives leading to the wrong act; for Hooker it is the right decision procedure leading to the wrong act. They hold that it is problematic for a moral theory to have this consequence. One way to formalize their complaint would be that global consequentialism is inconsistent with an intuitive principle about normativity:[158,159]

(1) If *s* ought to *X*, and *X*-ing leads *s* to *Y*, then it is permissible for *s* to *Y*.

However, the problem with this principle is that it is not only in conflict with global consequentialism, but with act-consequentialism, rule-consequentialism, many deontological theories and even common-sense morality. Consider the following case:

Pride after a fall

Penny is crossing a bridge when she hears a sudden splash and a frantic cry for help. She sees that a young boy has fallen into the river and that he cannot swim. Since no-

---

[155] Adams (1976), p. 470.

[156] Adams (1976), p. 479.

[157] Hooker (2004), § 5. See also Crisp (1992), Lang (2004).

[158] This principle is actually slightly weaker than what Adams would require, given that he was worried by merely *better* motives leading to wrong acts, and this principle only prohibits *right* motives leading to wrong acts. However, this is of little consequence as I shall soon show that even this principle is too strong.

[159] In this section we shall not need to distinguish between objective and subjective senses of normative and evaluative terms: the objection and my response to it work just as well regardless of which sense is used.

> one else is around, she wades in and rescues the boy, thereby saving his life. Later that day she is so proud of having saved someone's life that she decides to take the rest of the afternoon off, breaking a promise to help a friend move house.

In this case we can assume that Penny acted rightly in saving the boy and wrongly in breaking the promise to her friend. Moreover, it was this right act of saving the boy that led to the wrong act of breaking the promise. This is in contravention of (1). Since this type of case can arise in global consequentialism, act-consequentialism, rule-consequentialism, many deontological theories, and common-sense morality, none of these theories are compatible with (1).

It therefore seems that we must drop (1) on the grounds of it being overly strong. Adams and Hooker might claim that (1) was not sufficiently precise about the way that it formulates 'leads to', and that we could reformulate it with a more strict relationship between *X* and *Y*. For example:

(2) If *s* ought to *X*, and *s* cannot *X* without also *Y*-ing, then it is permissible for *s* to *Y*.

In this revised principle the required connection between *X* and *Y* is considerably tighter and thus the principle as a whole is considerably weaker. Indeed it is now too weak to apply to the commitment account of decision procedures, for it is possible to be committed to a decision procedure without performing all the acts that it recommends (through weakness of will, for example). However, (2) does still apply when we consider compliance with a decision procedure and perfect execution of a decision procedure. In Chapter 4 I showed that these accounts of decision making are less ethically relevant than the commitment account, but they are still legitimate focal points and global consequentialism will still produce judgments about them.

Must we accept (2)? No, for (2) is itself inconsistent with a widely accepted principle: actualism. Recall the case called *A predictable slump*, where Brian can go to the pub or stay at home.[160] If he stays home he can either work or slump in front of the television. If Brian's possible evenings were to be ranked in order of preference, they would be: working at home, drinking at the pub, watching television at home. The best evening thus involves Brian staying home, but he also knows that were he to stay home, he would in fact spend the evening slumping in front of the television.

As we have seen, opinion is divided on the question 'Ought Brian to stay at home?' Possibilists answer in the affirmative since Brian stays at home in the best future open to him. Actualists answer in the negative since were Brian to stay at home, this would be worse than the relevant alternative of going to the pub. They say that Brian ought not stay at home. They also accept that the answer to the question 'Ought Brian stay at home and work?' is 'Yes', because in this case the question is assessing the compound act of staying at home and working and the relevant alternatives are the courses of action over the same duration: staying at home and slumping in front of the television, or going to the pub and drinking.

---

[160] See section 3.3.2.

This actualist position is inconsistent with (2). Let *X* be *stay at home and work* and let *Y* be *stay at home*. Actualists accept that Brian ought to *X* and he cannot *X* without also *Y*-ing. However they deny that *Y* is permissible, since Brian ought not stay at home. (2) is thus incompatible with actualism: even actualist act-consequentialism.

In section 3.3.1, we examined theories of possibilist global consequentialism advanced by Crisp and Feldman. These theories are completely immune to the inconsistency objection as it was given by Adams and Hooker, for on these theories the right acts for you to perform are those acts you perform in the best world open to you, and the right decision procedure for you to follow is the decision procedure that you follow in the best world open to you. This possibilist formulation directly prohibits this type of conflict between the focal points and both Feldman and Crisp point to this as an advantage of their theories.[161]

The inconsistency objection thus appears to be an objection to all forms of actualism, with no particular connection to the assessment of multiple focal points. There are good reasons to accept actualism. Some of these were presented in section 3.3.1, and a much more detailed defence is has been provided by Jackson and Pargetter.[162] I will not go into the details of the actualism/possibilism debate here, but mention it merely to show that: *(i)* the form of inconsistency objection outlined by Adams and Hooker is not specific to global consequentialism, and *(ii)* there are many (act-)consequentialists who already deny that the objection is really problematic.

In summary, those people who ultimately find themselves persuaded to accept possibilism can accept a possibilist theory of global consequentialism like that of Crisp or Feldman. This will still allow one to apply consequentialism to every focal point and to answer the questions of action, decision making and motivation. Those who find actualism more plausible can accept an actualist theory of global consequentialism such as those I outline in sections 3.2.4 and 3.3.2. The only distinction between the actualist and possibilist forms of global consequentialism is in the set of answers they give to the actualist/possibilist puzzle cases such as that of *A predictable slump*. Because I find the actualist answer more plausible, that is the approach that I am assuming in this essay, but those who find the possibilist answer more plausible will see no loss in moving to a possibilist theory of global consequentialism.

6.1.2 Streumer

Another version of the inconsistency objection was given by Bart Streumer in his paper 'Can consequentialism cover everything?'[163] Streumer's argument is more explicit than those of Adams and Hooker, attempting to directly show that global consequentialist claims lead to a violation of the principle that 'ought' implies 'can'. Taking Parfit's case of Clare,[164] he sets out the argument like so:

---

[161] Feldman (1993), and Crisp (1992), pp. 143–5.

[162] Jackson and Pargetter (1986).

[163] Streumer (2003).

[164] See section 3.3.2 and Parfit (1984), pp. 31–40.

(1) Clare ought to love her child. *(premise)*

(2) Clare ought to benefit the stranger. *(premise)*

(3) Clare ought to both love her child and benefit the stranger. *(from 1 + 2) ???*

(4) Clare cannot both love her child and benefit the stranger. *(premise)*

We can then see that (3) and (4) together contradict the principle that 'ought' implies 'can'.

Violating the principle of 'ought' implies 'can' would indeed be a serious concern for global consequentialism; however as Campbell Brown has pointed out, Streumer's argument relies upon a dubious assumption.[165] Streumer cannot derive (3) from (1) and (2). He *could* derive the innocuous claim:

(3a) Clare ought to love her child and Clare ought to benefit the stranger

This new claim is a conjunction of two ought claims and can thus be derived from those two claims by logic alone. Claim (3), however, is an ought claim that applies to a conjunction. To get (3), we must go beyond mere logic and use a substantive principle such as the following:

Agglomeration

If *s* ought *X* and *s* ought *Y* then *s* ought *X* and *Y*.

The principle of agglomeration comes up frequently in discussions of deontic logic, and is controversial. In particular, it is rejected by actualists, for they hold that Brian ought to *stay home and work*, and Brian ought to *go to the pub*, yet they do not hold that Brian ought to *stay home and work and go to the pub*.

Streumer has responded to this point in a further paper.[166] In this he agrees that his former argument assumed agglomeration, but points out that he really only needed a weaker principle:

Combined 'oughts' imply 'can'

If *s* ought *X* and *s* ought *Y* then *s* can *X* and *Y*.

However, do we have sufficient reason to accept *combined 'oughts' imply 'can'*? Once again, this principle is rejected by actualists, for they hold that Brian ought to *stay home and work*, and Brian ought to *go to the pub*, yet they do not hold that Brian can *stay home and work and go to the pub*. This provides further evidence that the inconsistency objection is really an objection to actualism in all its forms and that it has nothing in particular to do with global consequentialism. As before, global consequentialists who are swayed by these criticisms of actualism can simply use a possibilist form of global consequentialism.

---

[165] Brown (2005).

[166] Streumer (2005).

Interestingly, Streumer puts forward his own possibilist version of global consequentialism:

> A thing is right if and only if it belongs to a combination of things that maximizes the good and that agents can bring about.[167]

This theory is much like those of Crisp and Feldman, except that it specifies rightness relative to a communal sense of 'can'. Instead of looking at what happens in the best world available to an individual agent, it defines rightness in terms of what happens in the best world that all agents together could realize. This has the advantage that it can talk about the right climate to have even if no individual can change the climate and the right person to elect even if no individual can change the outcome of the election.

However, it is also a version of collective consequentialism, not taking account of the fact that it might be disastrous for someone to perform a certain action if others do not play their part. For example, the best world that agents can bring about will not involve anyone stealing from anyone else and thus will not involve anyone wasting their time by locking their doors; however it is intuitively right to lock our doors since we know that some people won't act perfectly. In contrast, the role-based global consequentialism that I introduced in section 3.2.4 was set up to be able to assess rightness in both individual and collective senses. While Streumer's theory is interesting, those worried by the inconsistency objection would probably do better to take up Crisp's or Feldman's theories, or to construct a possibilist version of my role-based global consequentialism.

6.2 A regress concerning decision procedures

There is a potential regress when it comes to choosing which decision procedure to follow.[168] I want to act rightly and I look to consequentialism to find out what the right act is. However, consequentialism merely tells me that the right act is the one that makes the outcome best. This may be true, but I don't yet know which act that is. How then should I decide what to do? Consequentialism tells me that the right decision procedure to follow is the one that makes the outcome best. Once again, this may be true, but I don't yet know which one that is. How should I decide which decision procedure to follow? Global consequentialism tells me that I should decide this via the best available 'meta decision procedure'. This may be true, but I do not yet know which meta decision procedure this is…

This regress suggests that appeals to consequentialism do not tell us anything directly useful about how we are to act or to decide what to do. Even the decision procedure account that I have expounded at length appears to be just the second step along this infinite, and ultimately unenlightening, sequence.

---

[167] Streumer (2005), p. 229.

[168] Note that this regress is not one of those mentioned by Bales (1971). Something like it is, however, mentioned in Lang (2004), p. 227, and Smith (1991).

To some extent this is correct, but we must be careful not to exaggerate consequentialism's impotence. On its own, consequentialism does not give any concrete advice at any of these levels, but this is only because it is so utterly exposed to the facts. From the philosophical armchair, we commonly try to imagine ourselves in ignorance of the facts so that we might try to find *a priori* truths. In this state of ignorance, consequentialism will indeed give us no advice in a form that enables us to immediately recognise the right option. That is, all of its claims will be more like 'do that which promotes happiness' than 'do not lie'. For consequentialists, this is just the way that morality is and to say any more in ignorance of the facts would be impossible. Railton put this particularly well:

> 'A further objection is that the lack of any direct link between objective consequentialism and a particular mode of decision making leaves the view too vague to provide adequate guidance in practice. On the contrary, objective consequentialism sets a definite and distinctive criterion of right action, and it becomes an empirical question (though not an easy one) which modes of decision making should be employed and when. It would be a mistake for an objective consequentialist to attempt to tighten the connection between his criterion of rightness and any particular mode of decision making: someone who recommends a particular mode of decision making regardless of consequences would not be a hard-nosed non-evasive objective consequentialist, but a self-contradictory one.'[169]

I should note that it is only when making judgements about the objective rightness of acts or of decision procedures that consequentialism is open to the *facts*. When it makes judgments about what is right in the evidence-relative sense it is open to the *evidence* and for judgments about the belief-relative sense it is open to the agent's *beliefs*. Perhaps surprisingly, being open and responsive to any of these can produce a version of the regress.

The regress was created because we are considering how to make decisions in situations where the different ways of making the decision will have their own different costs and benefits. There is something quite strange about these situations. We know what to say about them in objective terms: the objectively right act is the one that will actually produce the best outcome, and the objectively right way to make decisions is the way that will actually produce the best outcome. We also know what to say in various subjective terms: the right act in the belief-relative sense is the one that will produce the expectably-best outcome, and the right way to make decisions in the belief-relative sense is the way that will produce the expectably-best outcome (taking the probabilities to be the agent's degrees of belief in each case). Similarly, we know what to say about the right options in the evidential sense.

The problem is that there still something lacking. There is at least one remaining sense that is even more 'internal' than any of these. It reflects the fact that the agent is either unable to determine what is right in any of the previous senses or would take so long to do so that it would be better for them not to attempt it. While the belief-relative and evidence-relative senses are meant to be internal, they do not quite capture the subtlety of this situation because they abstract away the very costs of deliberation which create the whole problem.

---

[169] Railton (1984), p. 156.

It is worth noting that when confronted with these situations, we do not actually succumb to a paralysis of decision making. Instead, we simply proceed with some form of imperfect decision procedure. Some of us are aware that we have sub-optimal decision procedures and make an occasional effort to think about how they can be improved. Sometimes these efforts lead to insights with which we modify our decision procedures, and at other times the additional introspection just serves to distract us. At no point do we spiral downwards to paralysis or upwards to a 'perfect' decision procedure that exactly balances the costs of complexity with the benefits of accuracy.

Can global consequentialism help us with our decision making in practice? Yes it can. It tells us the standard against which methods of decision making are to be judged, and this is useful information. We cannot use it to practically determine the perfect decision procedure in the objective, evidence-relative or belief-relative senses, but we can use it to make practical improvements to our existing decision procedures. For example, we can see that almost everyone would produce more good if their decision procedures had more stringent requirements for donations to the world's poorest people than they currently do. Too stringent a requirement may be counter-productive and finding the ideal level is indeed difficult, but in almost all current cases a small increase in their perceived requirement to aid would be beneficial.

We can also see that there is considerable merit in adopting a split level, self-modifying decision procedure like that advocated by Hare, for this approach would allow us to modify our decision procedures in light of new evidence or argument.[170] While this involves using a form of naïve calculation at the higher level, this is less alarming than it may first appear, because its application is limited by the split level structure. Its use is restricted to those times when we are not rushed, and since it is only used to choose principles rather than actions, there is less scope for our being tempted into personal bias. We can thus derive considerable benefit from applying consequentialism to decision procedures, even if we may not be able to derive the optimal benefit.

The regress, as I sketched it, is an extended version of the action-guidingness objection — that is, the claim that consequentialism is in some sense deficient as a moral theory because it does not tell us what we should do under immediately useful descriptions. The regress extends this to show that consequentialism cannot easily sidestep this with claims about the best decision procedure, since it also does not tell us which decision procedure this is under an immediately useful description. However, while consequentialism is of admittedly little use in finding the perfection it lauds, it is of significant benefit in finding *very good* acts, motives or decision procedures. It provides a standard against which to measure these things and this routinely helps us move from the good to the better, whether in acts, motives, decision procedures or many other objects of appraisal. As Mill put it:

[170] See section 4.1.2 and Hare (1981), Ch. 3; Hare (1989), Ch. 4 & Ch. 13.

> 'The corollaries from the principle of utility, like the precepts of every practical art, admit of indefinite improvement, and, in a progressive state of the human mind, their improvement is perpetually going on.'[171]

Finally, it must be said that consequentialism is not alone when it comes to this apparent regress. The regress occurs just as much in prudential decision making as it does in moral decision making. For example, if you are trying to finish a complex piece of writing for a deadline, there is a question about how much time should be spent on first order activities such as researching the literature or typing up the piece. To divide up the time available requires complex decision making. This decision making is itself an activity involved in the writing of the piece (a second order activity) and it is not clear how much time should be spent on it or in which way we should go about doing it. We could therefore spend some time deciding these matters, and so on. In practice, we muddle through such situations all the time, performing activities on the first one, two or three levels and typically leaving it at that. So long as we have a theory of prudence which is open to the facts (or open to the evidence, or open to the agent's beliefs), it will have to face this regress too.

These trade-offs also occur in other moral theories — or at least, they should. If moral theories are blind to the facts (or evidence, or beliefs) and prescribe one ultimate way of making decisions, then they are open to disaster. Suppose there were a demon who would cause unending misery if anyone failed to make their decisions according to a form of naïve calculation and that this were well known. If a moral theory does not have the flexibility to prescribe that we follow naïve calculation in such a case, it would knowingly embrace disaster. The only way to avoid this is to prescribe decision procedures based on their merits given the facts at hand, and theories that do so must face the same difficulties as the consequentialists.

In summary, there is a form of regress which affects consequentialists who distinguish between a criterion of rightness and a decision procedure. It makes it very difficult from an internal perspective to use consequentialism to work out the optimific acts or decision procedures, but it does not prevent consequentialism from providing very useful advice about which acts or decision procedures are likely to be better than others. Moreover these difficulties are by no means unique to consequentialism: they are shared by all theories of prudence and morality that do not knowingly embrace disaster.

6.3 The indistinctiveness objection

An old objection to the distinction between decision procedures and the criterion of rightness is that in making this distinction, consequentialism would cease to be: it would lose so much of its distinctive character that it would no longer be consequentialism at all. This objection has been made most famously by John Rawls and Bernard Williams. In *A Theory of Justice*, Rawls wrote:

---

[171] Mill (1861), Ch. 2.

> 'What we want to know is which conception of justice characterizes our considered judgments in reflective equilibrium and best serves as the public moral basis of society. Unless one maintains that this conception is given by the principle of utility, one is not a utilitarian.'[172]

In a footnote to this passage, Rawls illustrates it with the example of Richard Brandt's theory of ethics, wherein a society's public moral code is the one that maximizes utility, even if it does so without reference to a principle of utility.[173] On Rawls' view, such a theory cannot be counted as utilitarian. While Brandt's theory was intended as a form of rule-utilitarianism, Rawls would presumably extend his criticism to global consequentialism, claiming that it is not really consequentialism since it also does not require a maximising principle to play a part in the public moral basis of society (indeed global consequentialism need not even recommend that there *is* a public moral code, for it may be better if different groups have their own codes).

In *Utilitarianism: for and against*, Williams criticized sophisticated utilitarianism along similar lines:

> '…utilitarianism's fate is to usher itself from the scene … there seems to be nothing to stop, and a lot to encourage, a movement by which it retires to the totally transcendental standpoint from which all it demands is that the world should be ordered for the best, and that those dispositions and habits of thought should exist in the world which are for the best, leaving it entirely open whether those are themselves of a distinctively utilitarian kind or not. If utilitarianism indeed gets to this point, and determines nothing of how thought in the world is conducted, demanding merely that the way in which it is conducted must be for the best, then I hold that utilitarianism has disappeared, and that the residual position is not worth calling utilitarianism.'[174]

Both Rawls and Williams claim that if a theory recommends a principle of decision making that does not require one to attempt to maximize happiness, then the theory is not utilitarian. In this they are defining what is and is not a utilitarian theory, and their definition has the bizarre consequence that Mill and Sidgwick are not utilitarians, for we have seen that they both held that we should decide what to do in whichever way would be optimal.[175] Indeed, it is unclear whether any prominent philosophers who are commonly regarded as utilitarians would count as utilitarians under such a definition. It is true that many popular accounts of utilitarianism implicitly or explicitly suggest that utilitarianism requires what I call naïve calculation, but if the major figures in the field explicitly reject this, then surely we should take their word for what counts as utilitarianism. After all, it seems more plausible that Rawls and Williams were mistaken about what counts as utilitarianism than that all the major figures in the field were.

---

[172] Rawls (1971), p. 182.

[173] Brandt (1967), pp. 58–65.

[174] Williams (1973), pp. 134-135.

[175] Recall, for example, my quotations in section 1.3, or look directly in Mill (1861), Ch. 2 and Sidgwick

One of the key motivations behind the indistinctiveness objection is the idea that a moral theory is to be thought of as — by its very nature — a public set of principles for the running of a society. Rawls considered this a necessary condition for a theory of justice and called it the 'publicity condition'.[176] Under such a condition, the theories supported by Mill and Sidgwick about maximizing utility would not constitute 'moral theories'.

However, these theories do make many recommendations of a more public character. For example, they specify a best way of structuring society, a best decision procedure for each person's use and a best decision procedure for universal adoption. The first and third of these fit the publicity condition, and thus for proponents of the publicity condition these derived principles are perhaps more fitting of the title 'moral theory' than is the utilitarian doctrine itself. Proponents of the publicity condition could then see the core doctrine about maximizing happiness as some kind of higher level theory which is not itself a moral theory, but which provides an account of which moral theory is justified. Seen this way, it would be correct to say that the resulting 'moral theory' may be justified by utilitarianism but is not itself utilitarian.

This seems to me to be the most favourable interpretation of Rawls' and Williams' claims that the view does not deserve to be called 'utilitarian', for what is utilitarian is not on this account a moral theory, and what is a moral theory is merely justified by utilitarianism. However, I don't see this as much of a complaint against such sophisticated consequentialist theories. Advocates of these theories want to agree with Rawls and Williams on many points. For example, perhaps the best way to structure society would be under Rawls' two principles of justice: to a consequentialist this is an empirical matter and the two principles have much to recommend them. Likewise, the best decision procedures for individuals and the best decision procedure for universal use would have many properties that Williams would support, such as allowing for a certain amount of freedom in each agent's personal life.

Rather than dismissing the sophisticated view as not worthy of the name utilitarianism, nonconsequentialists should welcome it. They may not agree with the decision procedures or structures for society that consequentialism recommends, and they may not agree with the whole program of justifying such decision procedures or structures in terms of the utility they create, but they will surely agree that it is much better than the unsophisticated support of naïve calculation that often goes by the name 'utilitarianism'. Moreover, they should welcome the sophisticated consequentialists as allies against these misguided advocates of naïve calculation.

Still, consequentialists may want to ask themselves whether they have lost their distinctive position in the move to sophisticated approaches such as global consequentialism. To this question, I think the answer must be 'No'. Contrary to Williams' claim that consequentialism 'determines nothing of how thought in the world is conducted', it would actually provide very tight constraints on how thought is to be conducted, requiring that it be conducted in the way that leads to the best possible outcome. This is distinctively consequentialist if

---

(1907), pp. 489–90.

[176] Rawls (1971), p. 133.

anything is.[177] It is true that the sophisticated theory would not endorse many attempts to naïvely calculate what would be best, but why were consequentialists ever interested in such calculation? Only because it was thought to be a useful method for arriving at good outcomes: because it was perceived to be an *instrumentally* good decision procedure. To support it even when it is known to lead to worse outcomes would be 'rule-worship' and ultimately self-defeating.

Have sophisticated consequentialists lost sight of their objective? No. A golfer with a known slice (a disposition to hit the ball to the right of where one aims) can consciously aim to the left of the hole and thereby better fulfil her goal. The goal was to get the ball in the hole and aiming directly at the hole was merely instrumental to this goal. We can tell that the golfer's ultimate goal is to get near the hole even though she does not aim directly at it by noting the counterfactual that if she were to develop a hook, she would then start to aim to the right of the hole.[178] Alternatively, we could determine her ultimate goal by simply asking her.

Similarly, sophisticated consequentialists may find that in 'aiming thus at something else, they find happiness by the way.'[179] We can tell that it is happiness (or the good) that is their true goal by noting that if some other method were known to be better at producing the good, then they would switch to that method. Finally, as with the golfer, we could also just ask them, for people are sometimes consciously aware of when their ultimate goal differs from the sub-goal at hand.

Let us conclude with another quotation by Mill, which ably demonstrates just how clearly this idea has been understood in the consequentialist literature since its very conception:

> 'I do not mean to assert that the promotion of happiness should be itself the end of all actions, or even of all rules of action. It is the justification, and ought to be the controller, of all ends, but is not itself the sole end.'[180]

6.4 Blame and guilt

A final objection to global consequentialism is that it makes incorrect or puzzling claims about attitudes such as blame and guilt. As Hooker writes:

> 'Global consequentialism claims that you should be blamed if and only if blaming you will produce the best consequences, and that you should feel guilty if and only if this will produce the best consequences. Suppose that for some reason the best consequences would result from blaming you for following the prescribed decision procedure (and thus doing *x*). But surely it is paradoxical for a moral theory to call for you to be blamed although you followed the moral decision procedure mandated by

[177] This point has been made very well by Parfit (1984), p. 42.

[178] A similar point is made by Railton (1984), p. 145.

[179] Mill (1873), Ch. 5.

[180] Mill (1843), Bk. VI, Ch. 12, §7.

the theory.'[181]

Thus, while common sense would suggest that a person should be blamed if and only if they have done something (subjectively) wrong, global consequentialism assesses blame just like any other focal point and looks only at the outcome of blaming someone. Similarly, common sense would suggest that a person should feel guilty if (and only if) they have done something wrong, but global consequentialism looks only at the outcome of feeling guilty.

To some extent this objection is not specific to global consequentialism. Similar remarks have long been made about act-consequentialism requiring us to blame people whenever that will improve the outcome.[182] However, act-consequentialism assesses only *acts* of blaming. Global consequentialism can go one step further and assess someone's holding an attitude of blame towards someone else. This objection is thus similar to the old objection, but slightly more serious.

The first thing that global consequentialists can say in response is that the objection only works if we can choose whether to blame people or whether to feel guilty. If we cannot, then, since 'ought' implies 'can', global consequentialism will not say that you should blame a person or feel guilty. When blaming someone would produce a good outcome, global consequentialism would still say that it is better to blame them than to not blame them, but its claims would be merely evaluative rather than normative and consequently less troubling. However, for the sake of argument let us grant that we do have a degree of choice in such matters and see what this would mean for global consequentialism.

Those who criticize global consequentialism for its assessment of blame and guilt are presumably looking for a simple connection between these attitudes and wrongness: we should blame people when they act wrongly and we should feel guilty when we act wrongly. However, there is a large price to pay for such a simple connection. The problem is that it is completely insensitive to the consequences of blame and guilt.

To make this clear, suppose that an evil demon will produce unending misery unless you hold an attitude of blame towards your mother (who has not done anything to deserve this blame). Surely in such a case you should do everything in your power to come to hold this attitude of blame. A moral theory need not say that you should blame your mother (in the sense of holding an attitude of blame towards her), for the theory does not have to provide any assessment at all of attitudes. However, there are still the acts involved in coming to blame your mother and any plausible theory must say that you should try exceedingly hard to make yourself blame her. Such theories are thus forced to say things that are very similar to those of global consequentialism. These theories may even be more 'schizophrenic' if they retain their commonsense claims that you should blame someone if and only if they have acted wrongly, for it would then be the case that you should not blame your mother and you should try very hard to make yourself blame your mother. A similar story can, of course, be told for feeling guilty.

---

[181] Hooker (2004).

[182] See, for example, Sidgwick (1907), p. 428, and Smart (1973), pp. 49-55.

In summary, this objection is not damaging for global consequentialism. Either we cannot choose whether to blame someone or to feel guilty, in which case global consequentialism does not normatively assess blame or guilt, or we can choose, in which case all plausible theories have to say very similar things to global consequentialism.

Chapter summary

In this chapter I have defended global consequentialism against four objections. I first showed that the inconsistency objection is really an objection to actualism and that its only effect on global consequentialism is to push some people towards possibilist theories of global consequentialism. I then explored a potential regress regarding decision making, showing firstly that it does not prevent the decision procedure approach from providing useful advice, and secondly that it cannot have any particular force against global consequentialism since all plausible theories of prudence and ethics are also subject to the same regress.

The third objection was that a theory which applies the consequentialist criterion to motives and decision procedures is no longer really consequentialist. I showed that this objection is misguided as this is the form of consequentialism that most famous consequentialists have held and it *does* still make distinctively consequentialist claims. Finally, I looked at the objection that global consequentialism gives an inappropriate assessment of blame and guilt, showing that global consequentialism would only provide normative assessment of these if we have a degree of choice about when we blame or feel guilty, but that in this case all ethical theories that do not knowingly embrace disaster must make similar claims.

Chapter 7

Conclusions

Normative ethics is often divided into three great traditions: consequentialism, deontology, and virtue ethics. Each captures a strong intuition about the nature of ethics: that what is ultimately important is that we produce the best possible outcome, that ethics is a system of moral rules for governing our behaviour, and that ethics is about living a life that instantiates the virtues. When each of these intuitions is considered on its own, it seems very persuasive, but the traditions they give rise to are thought to be in fundamental disagreement with each other. This conflict raises questions about the possibility of real progress in ethics: there appear to be irreconcilable differences at the most fundamental level and very little sign of convergence in moral views.

Global consequentialism offers a different perspective. It is a solidly consequentialist framework, but one that is in harmony with many of the intuitions behind deontology and virtue ethics. Global consequentialism agrees with deontology that intuitive moral rules such as 'do not kill' or 'keep your promises' are central to ethics, and it largely agrees about the content of such rules. It recognises the importance of being deeply committed to these rules in a way that excites moral compunction. It accepts that such rules need to allow space for personal projects, and that they will be agent-relative: having a special place for close family and friends. It denounces naïve calculation as a deeply flawed method for making moral decisions.

Global consequentialism agrees with virtue ethics that living a life of virtue is central to ethics. It acknowledges that a person's character plays a pivotal role in shaping their life and greatly affects the lives of those around them. Moreover, it largely agrees with virtue ethics about which character traits are good or bad. It assesses agents and not just acts, taking seriously questions such as 'How should I live?' and 'What kind of person should I be?' Global consequentialism is also sympathetic to the idea that practical wisdom is of significant moral importance, for the best decision procedures would have to leave room for intuitive judgments, and such judgments would benefit immensely from the application of hard-won practical wisdom.

Of course, there are numerous varieties of deontology and virtue ethics, and there will be many claims that global consequentialism cannot accommodate. For example, some virtue ethicists hold that the reason someone should be virtuous is not for its benefits to others, but because it is part of one's own *eudaimonia*. Also, virtue ethics is sometimes associated with

moral particularism or the 'uncodifiability of ethics'. Global consequentialism can accommodate this view if it is understood at the level of decision procedures, since the best decision procedures will not be able to codify everything and will require significant room for the agent's judgment. It can also accommodate it if understood as the epistemic claim that it is very difficult to determine what is right and that there are no readily usable principles through which someone can determine what they objectively ought to do. However, global consequentialism does claim that at the most fundamental level ethics is very structured and principled and can in this sense be completely codified.

An important conflict between global consequentialism and deontology concerns the question of which acts are right. While there is some potential for agreement about the right rules, deontology indirectly assesses acts via these rules and thus arrives at a very different conclusion about which acts are right. On the surface, this is a dramatic disagreement between the theories, but it is lessened when we consider what rightness and wrongness entail in each of the theories. For deontologists, a wrong act is not merely suboptimal but is beyond the pale, warranting guilt, blame, and sanction. When a global consequentialist calls an act wrong, this does not necessarily entail any of these things. Indeed, a deontologist's use of the term 'a wrong act' is much more closely analogous to what a global consequentialist would call 'an act that does not comply with the decision procedure it would be (subjectively) best for the agent to accept'. The difference is thus partly one of terminology: both theories see problems with acts that lead to bad consequences and both theories see problems with performing acts that are not in accordance with the right rules or decision procedures, but they disagree on which group is appropriately termed 'wrong'. However, this is still a point of contention between the theories and no doubt some would count this as a reason to prefer deontology.

There are also other conflicts between the two approaches which the move to global consequentialism can only partly resolve, such as what we should say about good acts done from bad intentions. I do not pretend that global consequentialism captures *everything* that leads people to deontology, but it does capture a great deal about the essential role of rules and decision procedures in moral life and this lessens the gap between consequentialism and deontology, just as its consideration of the importance of motivation and character lessens the gap between consequentialism and virtue ethics.

This increased harmony between global versions of consequentialism and other approaches to ethics was grasped very keenly by Sidgwick. At the end of his life he wrote a series of notes on the development of his thought on utilitarianism, and these notes were published posthumously as part of the preface to the sixth edition of *The Methods of Ethics*. In them he wrote:

> '… I tried to say what I had found: that the opposition between Utilitarianism and Intuitionism was due to a misunderstanding. … I could find no real opposition between Intuitionism and Utilitarianism. … the best examination I could make of the Morality of Common Sense showed me no clear and self-evident principles except

> such as were perfectly consistent with Utilitarianism.'[183]

It was also understood by Mill, when he wrote:

> 'Those who adopt utility as a standard can seldom apply it truly except through the secondary principles; those who reject it, generally do no more than erect those secondary principles into first principles. It is when two or more of the secondary principles conflict, that a direct appeal to some first principle becomes necessary … for the sake of the systematic unity and coherency of ethical philosophy.'[184]

He thus saw that one of the central roles of the greatest happiness principle was to justify a system of rules or principles that one could readily apply in order to guide one's action. He recognized that these very rules were often seen as fundamental principles in deontological theories or in common-sense morality. He thus did not see the relationship between the consequentialist's criterion of rightness and a deontologist's rules as one of competition for the same role, but rather that the former is a way of justifying and unifying the latter. His remarks upon the importance of character showed that he held the same view there: 'I hold that the very question, what constitutes this elevation of character, is itself to be decided by a reference to happiness as the standard.'[185] Much of the apparent conflict between these approaches to ethics has thus consisted in them 'talking past' each other, rather than as direct competition for a single role in our moral thought. Moreover, this was observed more than a century ago by those who first took consequentialism beyond action and applied it to decision making and motivation.

This essay has been an attempt to explore and strengthen this approach. I showed that it could not be adequately achieved with (mere) act-consequentialism, or with rule- or motive-consequentialism. I therefore explored the various forms of global consequentialism, demonstrating the need to assess each focal point within a particular role, and providing two complete global consequentialist theories, either of which would allow us to coherently and precisely apply consequentialism to decision making and motivation. I then examined many ways in which we could interpret the questions of decision making and motivation, showing which interpretations are most central to our ethical thought. Finally, I examined and resolved four objections to global consequentialism.

Since a picture sometimes leads to better consequences than a thousand words, I shall conclude with a diagram. The figure below represents the relationships between focal points in deontology, virtue ethics and several versions of consequentialism. It therefore provides a rough summary of the structures of the different theories which I have discussed and their relationships to one another. The arrows represent the hierarchy of justification within each theory, and can be roughly read as 'justifies'.[186]

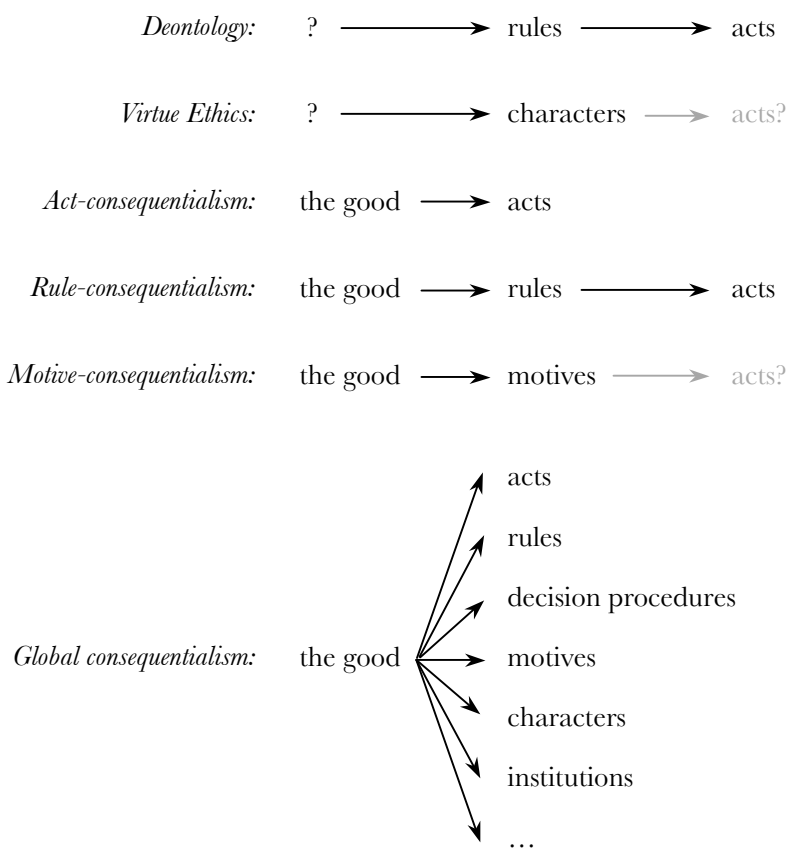
---

[183] Sidgwick (1907), pp. xxii–xxiii.

[184] Mill (1838), p. 111.

[185] Mill (1843), Bk. VI, Ch. 12, §7.

[186] This figure, of course, owes a great deal to the figures in Kagan (2000). My contribution has been to



extend it to provide a sketch of the structure of deontology and virtue ethics along similar lines. I have shown the focal point of 'acts' in grey for virtue ethics and motive-consequentialism, because some forms of these theories assess acts and some do not.

Appendix

Right or Best?

This appendix analyses a question that arose in section 3.2.3: should global consequentialism apply normative terms (should, right, ought…) to all focal points (*normative global consequentialism*), or should it restrict the use of these terms to acts (*semi-normative global consequentialism*). I have already considered the possibility that the distinction might be purely terminological, so here I will consider it as a substantive dispute.

There are three major concerns for normative global consequentialism. Firstly, there is the question of whether it is even conceptually possible to apply normative terms to non-acts, then there is the question of who 'owns' the resulting normative claims, and finally there is the question of how normative claims concerning non-acts relate to the principle that 'ought' implies 'can'.

There are also three concerns for semi-normative global consequentialism. First there is the question of whether this involves unnecessarily privileging acts, then the question of whether it is really consequentialist, and finally the question of whether we are left with sufficient resources to answer the question of decision making.

Concerns for normative global consequentialism

One reason to hold the asymmetry between acts and other focal points is if one simply does not believe that normative terms *can* be applied to things other than acts. This belief that acts are the exclusive (or at least primary) bearers of rightness, or obligatoriness, is widespread. Many philosophers believe that it doesn't make sense to talk of the right motives to have, or the right decision procedure to follow. Instead they would see it as some kind of category mistake.

However, there is at least some room for disagreement here since Pettit and Smith repeatedly apply (moral) rightness to things other than acts and clearly don't think that what they say is nonsense. It is possible that in their quest for a universal consequentialist theory, they simply confused themselves into committing a category mistake, but this seems less likely when we consider the way in which normative terms are used in everyday conversation. For example,

consider the following 'ought' claims:[187]

(1) You ought to be there on Sunday.

(2) You ought be at work by nine.

(3) You ought to be more considerate.

(4) We ought to be ashamed of ourselves.

(5) There ought to be a law to deal with this.

(6) This is the way things ought to be.

(7) She ought to know her times tables by now

(8) You ought to believe in God

(9) Mothers ought to have equal rights

These are quite typical uses of the word 'ought' and are all frequently uttered. They are all grammatically correct and they all have fairly clear meanings.[188] It is notable that many of them are in the form 'ought to be'. This recalls a distinction in normative thinking (particularly in deontic logic) between 'ought to do' claims and 'ought to be' claims. It is said that these are the two key types of normative claims and that the 'ought to do' claims are unique in attaching to a specific agent (the agent whose performance of the act is in question), while the 'ought to be' claims are more like claims of goodness, expressing an ideal state of affairs without placing demands on any particular person to bring it about.[189]

One might try to fashion a form of global consequentialism around this distinction, using 'ought to do' for acts and 'ought to be' for other focal points.[190] However, I do not think that we are forced into such a move, for I doubt that this distinction really captures what is going on here. Of the above claims, only (5) and (6) seem to be typical of this account of 'ought to be'. On claims (1)–(3) there is a clearly specified agent who appears to 'own' these obligations, despite their being phrased as 'ought to be'. Claim (4) is similar, though it applies to more than one agent. Defenders of the distinction might argue that these claims are actually disguised 'ought to do' claims, which are merely expressed as 'ought to be' claims.

---

[187] I focus on 'ought' for the next few pages because the literature on the related aspects of normativity is all couched in terms of 'ought': 'ought to be' versus 'ought to do'; ownership of oughts; 'ought' implies 'can'. I believe that it could also be carried out in terms of 'should' or 'right'.

[188] Insofar as any moral claims have clear meanings.

[189] See, for example, Castañeda (1970).

[190] For the sake of consistency, it might even be better to forsake 'ought to do' and describe acts too using 'ought to be'. For example: 'it ought to be that Michael walks to the shop' or 'Michael ought to be walking to the shop between 12:00 pm and 12:10 pm'.

However, I do not find this convincing, as they do not refer to specific actions and, especially in claims (3) and (4), I do not see how they could be reworded to do so. Finally, in claims (7)–(9) the phrase 'ought to be' does not even occur.

Instead, I offer the following explanation of what is going on. To produce a grammatical sentence, we must follow the term 'ought to' with a verb.[191] Most verbs describe actions and are known as 'action verbs'. Some do not, and these are called 'state verbs'. An especially versatile state verb is 'be', but there are others too. Examples include: 'have', 'possess', 'believe', 'know', 'cope', 'flourish', 'fail', 'succeed', 'require', 'fit', 'resemble', 'love', 'understand', 'need', 'want', 'prefer', 'remember', 'forget' and 'own'. Many of these are used in everyday ought claims.

We might thus explain the common assumption that 'right' and 'ought' can only apply to actions by the simple fact that most verbs are action verbs. We might also explain the reliance on 'ought to be' as the only alternative in terms of 'be' being the most versatile and prevalent of the state verbs. Indeed, with the heavy pedagogical focus on action verbs (through the common refrain of 'verbs are *doing* words'), it is understandable that many English speakers are only dimly aware that there *are* state verbs.

If we therefore grant the existence of obligations involving state verbs like 'have' and 'possess', we can make normative global consequentialist claims such as 'Oxford ought to have had a temperate climate in the 18th Century' and 'Alice ought to possess such and such a motive-set' or, alternatively, 'Alice ought to be motivated by such and such a motive-set'.

We might also wonder how this talk of 'ought' connects with 'right'. Philosophers typically consider the terms to be inter-translatable and that it is a matter of convenience as to which term is given prominence in an ethical theory. The standard consequentialist connection between the two is that (excluding ties) you ought to do something if and only if it is right. In the case of a normative tie between x and y, they are both right, but neither is obligatory. It seems natural to extend this connection to things other than actions. For example, Oxford ought to have had a temperate climate in the 18th Century if and only if temperate is the right climate for Oxford to have had in the 18th Century.

However, there is one catch. Suppose that of all the people Maxwell could kill, the consequences would be least bad if he killed Helen, but this would still be much worse than if he were to kill no-one. We would then say that the best person for Maxwell to kill is Helen and, on Pettit and Smith's theory, that the right person for Maxwell to kill is Helen. However, it doesn't follow that Maxwell ought to kill Helen. The problem arises because it is possible for Maxwell to kill no-one at all and this would lead to the best outcome. This situation cannot arise with acts, because (on consequentialist accounts of acts at least) we must always perform some act, even if it is sitting still for a moment.

---

[191] Note that I am not demanding that our ought claims *be* grammatical (we shall later see that John Broome has very good reasons for rejecting this), but that the underlying grammar helps explain why philosophers have focused attention on particular groups of ought claims.

We can resolve the problem in two different ways. The first is to keep the simple connection between best and right, as Pettit and Smith do, but to give a more complicated connection between right and ought. We can say that (excluding ties):

x ought to be in role *R* *iff* *x* is the right *f* to be in role *R*
and it isn't better for no *f* to be in *R* than for *x* to be in *R*

Alternatively, we can keep the simple connection between ought and right, but use a more complicated connection between right and best:

x is the right *f* to be in role *R* *iff* *x* is the best *f* to be in role *R*
and it isn't better for no *f* to be in *R* than for *x* to be in *R*

Through either of these translations, the above defence of normative assessment of focal points other than acts could be seen as applying to claims about rightness as well as claims about ought.

Let us now return to the question that has arisen concerning the ownership of ought claims. When we say 'the right thing for Maria to do is to donate the money' or 'Maria ought to donate the money', we think of these claims as being 'owned' by Maria: they exert a normative pull upon her specifically. Maria is the bearer of the obligation to donate the money. Perhaps others should try to make her do so, perhaps not. That depends on their obligations and not solely upon her own. In contrast, the claims that 'the right climate for Oxford to have in the 23rd Century is temperate', or that 'Oxford ought to have a temperate climate in the 23rd Century' don't seem to be directly owned by anyone. Perhaps we can all play our part in making sure that Oxford maintains a temperate climate into the 23rd Century, but the obligation does not apply to an individual and it is not clear that it applies to any specific group of people, unless that includes absolutely everyone.

For some claims it is ambiguous as to whether they are owned by anyone. For example, if we claim that 'the right set of beliefs for Mary to hold is such-and-such', does this express an obligation that is owned by Mary? Perhaps its ownership is shared by others who play a significant role in shaping her beliefs, such as her teachers and parents. They might be able to control whether she believes that 12 times 12 is 144 much better than she herself can.

We have seen that the distinction between 'ought to do' and 'ought to be' is thought to hold the answer here, with all 'ought to do' claims expressing ownership and all 'ought to be' claims being unowned. However, we have also seen that this account does not seem to account for all common ought sentences. John Broome provides a different approach.[192] He denies the standard division and holds that we can separately specify both the obligation and its owner. One of his examples concerns Julie, who ought to have more work to do. It makes perfect sense to consider that Julie owns this obligation and it also makes sense to think that it might instead be her boss's responsibility and thus that her boss owns this obligation. Broome's system allows us to express each of these possibilities explicitly.

[192] Broome (2008).

Broome's system is also of interest to the global consequentialist for its focus on normatively assessing arbitrary propositions. These propositions may involve actions, such as 'Maria donates the money', but they could also express thoughts that don't involve actions, such as 'Julie has more work'. While Broome's system is not committed to consequentialism, it *is* global. One might think of it as a theory of *global normativity*, and we can then see Pettit and Smith's theory as lying within the intersection of global consequentialism and global normativity.

Ought claims in Broome's system are expressed in the form:

S ought that *F*

Where *S* is the owner and *F* is the proposition. For instance:

Maria ought that Maria donates the money.

Julie ought that Julie has more work.

The boss ought that Julie has more work.

One complaint about these constructions is that they are ungrammatical. However, Broome provides a convincing account of how: *(i)* we cannot say such things unambiguously within the existing grammar of 'ought' and *(ii)* we nevertheless possess the concepts involved. He sees his expressions of the form '*S* ought that *F*' as a grammatical innovation, that extends the grammar of 'ought' and overcomes a mere quirk of the English language which was preventing us from verbalising such thoughts previously.[193] Moreover, the focus on arbitrary propositions rather than actions allows for more complex obligations, such as obligations to conditional statements, conjunctions and disjunctions.

There are thus several options for how normative global consequentialists can account for the ownership of oughts. They can let the grammar define the ownership, they can have all oughts be unowned, they can have all non-act oughts be unowned, or they can specify which oughts are owned and by whom. This last route allows more flexibility, but requires an explanation of how oughts come to be attached to specific people. We shall examine such a theory in section 3.2.4.

The final issue for the normative global consequentialists is that of possibility. If we accept that 'ought' implies 'can', then there must be some sense in which the obligations of global consequentialism can be met. Since Pettit and Smith do not explicitly account for such possibility in their formula, it may well produce obligations that are impossible to fulfil, such as the right act being to fly away, when the agent cannot fly.[194] The obvious way to remedy this is by imposing some appropriate form of possibility upon the focal point. For example,

---

[193] He also points out that if we are really averse to modifying the existing grammar, we can just make a simple formal language to express the same ideas.

[194] Pettit and Smith get very close to this when they say that global consequentialism determines 'whether it is right for it to be rainy or cloudy or sunny' [Pettit and Smith (2000), p. 122].

Parfit talks of the best possible motives and climates, and specifies this as causal possibility. Pettit and Smith could follow suit by inserting the world 'possible' after the word 'best', as in: 'the right *x* is the best *possible x*'. A full account would then require an analysis of the appropriate type of possibility, and this would presumably involve the owner of the obligation (if any). However, since there is significant debate about the nature of such possibility, let us just say that we can use whichever type of possibility it is that ends up being involved in 'ought' implies 'can'.

Concerns for semi-normative global consequentialism

Now we come to the concerns for semi-normative global consequentialism. The first of these is that it privileges acts, by singling them out for normative treatment. This criticism holds only if it is possible to normatively assess things other than acts. If this is false (as most semi-normative global consequentialists presumably believe), then they are not really privileging acts but are merely assessing each thing in all ways that it can be morally assessed.

The second concern is that semi-normative global consequentialism is not really consequentialist, for it is quite compatible to say that '*x* is the best set of motives for Jane to have', '*x* is a better set of motives for Jane to have than *y*' and yet '*y* is the right set of motives for Jane to have'. This mirrors a pair of claims that a nonconsequentialist could make about acts: '*z* is the best act', but '*z* is the wrong act'.[195] Perhaps the semi-normative global consequentialist is agreeing that Jane's possession of motives *x* would lead to more good, but thinks that motives *y* are more godly, and thus right. Again, this criticism only holds if it is possible to have normative assessment of things other than acts. If not, then the nonconsequentialist possibility is no longer assertable, and the problem vanishes.

The final concern is that semi-normative global consequentialism will not answer the question of decision making or the question of motivation. We ask 'How should I decide what to do?' and 'How should I be motivated?', but semi-normative global consequentialism only answers that 'the best decision procedure for you to follow is *x*' and 'the best set of motives is *y*'. These are merely evaluative answers to the normative questions. However, once again, if there are no valid normative assessments of things other than acts, then there isn't anything more that a theory *could* offer. The questions themselves would be misguided, and the answers would be providing the only possible kind of moral assessment of decision procedures and motives. In this case, it seems that we would be best to think of the questions 'What is the best way for me to decide what to do?' and 'What is the best way for me to be motivated?' as the appropriate formulations of the questions of decision making and motivation, in which case semi-normative global consequentialism can plausibly provide the answers.

---

[195] It is not clear how many nonconsequentialists actually endorse something like this. Nonconsequentialists believe that it is possible for the action that leads to the best consequences to be wrong, but they don't need to believe that the *best* action is wrong since they don't have to define the goodness of actions in terms of their consequences.

In conclusion, it seems that there is an open question as to whether we can normatively assess things other than acts. If we can't, then semi-normative global consequentialism appears to be able to provides a satisfactory answer to the questions of decision making and motivation, and normative global consequentialism is impossible. If we can, then we should use a version of normative global consequentialism. In this case, we must decide on an appropriate theory of possibility to take account of how 'ought' implies 'can' and we must also decide how the resulting obligations are owned, if at all.

Railton's approach

Finally, let us return to Railton's theory of global consequentialism which goes significantly further in the direction of mere evaluation. Firstly, he abandons normative terms altogether in the consequentialist part of his theory, arguing that the consequentialist account of wrongness as suboptimality is too much at odds with the pre-philosophic conception of wrongness as truly unacceptable behaviour. Thus, while he proposes holding a nonconsequentialist account of normative terms alongside his global consequentialist theory, the theory itself does not speak of rightness at all.

In the context of act-consequentialism this jettisoning of normative terms has been called *scalar consequentialism* and has been advanced by Frances Howard-Snyder and Alastair Norcross.[196] We could thus use the term *scalar global consequentialism* to refer to theories which evaluate everything but apply no normative terms. They would have the advantage of treating acts and non-acts exactly alike without needing to apply normative terms to the non-acts. We could then summarize these theories as follows:[197]

	Acts	**Rules**	**Motives**	**Everything else**
Act-consequentialism	Rightness	-	-	-
Rule-consequentialism	Rightness	?	-	-
Motive-consequentialism	-	-	Rightness	-
Semi-normative global consequentialism	Rightness	Betterness	Betterness	Betterness
Normative global consequentialism	Rightness	Rightness	Rightness	Rightness
Scalar global consequentialism	Betterness	Betterness	Betterness	Betterness

Table 1.

[196] Howard-Snyder and Norcross (1993).

[197] I have used a hyphen to represent that the theory says nothing about a given focal point and a question mark to indicate that it offers some kind of assessment, but whether it is rightness or betterness or something else depends on the particular version of the theory. Note that I have used Adams' version of motive-consequentialism [Adams (1976)]. Some other authors assume motive-consequentialism to be analogous to rule-consequentialism and thus say that it indirectly specifies the right acts in terms of the

However, Railton goes even further than this, abandoning the most common types of evaluation as well. He does not speak of the goodness or the moral value of focal points on the grounds that these terms will also be too out of touch with the pre-theoretic concepts. Instead he assesses focal points as being 'better rather than worse from a moral point of view' and uses the term 'more morally fortunate'. Unfortunately, it would take us too far off track to explore these aspects of Railton's theory, but it is important to note the possibility of global consequentialist theories that abandon rightness and even goodness.

---

right or best motives.

References

Adams, R. M. 1976. 'Motive utilitarianism', *Journal of Philosophy* 73:467–481.

Austin, J. 1832. *The province of jurisprudence determined*, H. L. A. Hart (ed.), (London: Weidenfield), 1954.

Bales, R. E. 1971. 'Act-utilitarianism: account of right-making characteristics or decision making procedure?', *American Philosophical Quarterly*, 8:257–265.

Bentham, J. 1789. *An introduction to the principles of morals and legislation*, (New York: Hafner), 1961.

Braddon-Mitchell, D. 2004, 'Freedom and binding consequentialism', unpublished.

Brandt, R. 1963. 'Toward a credible form of utilitarianism', in H.-N. Castañeda and G. Nakhnikian (eds.), *Morality and the language of conduct*, (Detroit: Wayne State University Press), 107–43.

Brandt, R. 1967. 'Some merits of one form of rule utilitarianism', *University of Colorado Studies in Philosophy*, 3:39–65.

Brink, D. O. 1986. 'Utilitarian morality and the personal point of view', *Journal of Philosophy* 83:417–38.

Broome, J. 2008. Unpublished manuscript for a book on reasoning and rationality.

Brown, C. 2005. 'Blameless wrongdoing and agglomeration: a response to Streumer', *Utilitas* 17:222–5.

Carlson, E. 1995. *Consequentialism reconsidered*, (Dordrecht: Kluwer).

Castañeda, H-N. 1970. 'On the semantics of the ought-to-do', *Synthese* 21:449–68.

Crisp, R. 1992. 'Utilitarianism and the life of virtue', *Philosophical Quarterly* 42:139–160.

Driver, J. 2001. *Uneasy virtue*, (Cambridge: Cambridge University Press).

Feldman, F. 1975. 'World utilitarianism', in Keith Lehrer (ed.), *Analysis and Metaphysics*, (Dordrecht: Reidel), 255–271.

Feldman, F. 1986. *Doing the best we can*, (Dordrecht: Reidel).

Feldman, F. 1993. 'On the consistency of act- and motive-utilitarianism: a reply to Robert Adams', *Philosophical Studies* 70:201–212.

Gibbard, A. 1990. *Wise choices, apt feelings*, (Oxford: Oxford University Press).

Gruzalski, B. 1986. 'Parfit's impact on utilitarianism', *Ethics* 96:760–783.

Hare, R. M. 1981. *Moral thinking: its levels, method and point*, (Oxford: Clarendon Press).

Hare, R. M. 1989. *Essays in ethical theory*, (Oxford: Oxford University Press).

Hooker, B. 2000. *Ideal code, real world: a rule-consequentialist theory of morality*, (Oxford: Oxford University Press).

Hooker, B. 2004. 'Rule-consequentialism', in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2004 Edition)*, URL = <http://plato.stanford.edu/archives/spr2004/entries/consequentialism-rule/>.

Howard-Snyder, F. and Norcross, A. 1993. 'A consequentialist case for rejecting the right', *Journal of Philosophical Research* 18:109–25.

Jackson, F. and Pargetter, R. 1986. 'Oughts, options, and actualism', *Philosophical Review* 95:223–255.

Jeffrey, R. C. 1990. *The logic of decision*, 2nd edition, (Chicago: University of Chicago Press).

Kagan, S. 2000. 'Evaluative focal points', in B. Hooker, E. Mason and D. E. Miller (eds.), *Morality, Rules and Consequences: A Critical Reader*, (Edinburgh: Edinburgh University Press), pp. 134–155.

Kavka, G. 1978. 'Some paradoxes of deterrence', *Journal of Philosophy* 75:285–302.

Lang, G. 2004. 'A dilemma for objective act-utilitarianism', *Politics, Philosophy and Economics*, 3:221–39.

Louise, J. 2006. 'Right motive, wrong action: direct consequentialism and evaluative conflict', *Ethical Theory and Moral Practice* 9:65–85.

Lyons, D. 1965. *The forms and limits of utilitarianism*, (Oxford: Oxford University Press).

Mason, E. 2002. 'Against blameless wrongdoing', *Ethical Theory and Moral Practice* 5:287–303.

McCloskey, 1963. 'A note on utilitarian punishment', *Mind* 72:599.

Mill, J. S. 1838. 'Bentham', in John M. Robson (ed.) *Collected Works (vol. X)*, (Toronto: University of Toronto Press), 1985, pp. 75–116.

Mill, J. S. 1843. 'A system of logic: ratiocinative and inductive', in John M. Robson (ed.) *Collected Works (vol. VIII)*, (Toronto: University of Toronto Press), 1974, pp. 1–952.

Mill, J. S. 1861. 'Utilitarianism', in John M. Robson (ed.) *Collected Works (vol. X)*, (Toronto: University of Toronto Press), 1985, pp. 203–260.

Mill, J. S. 1873. 'Autobiography', in John M. Robson and Jack Stillinger (ed.) *Collected Works (vol. I)*, (Toronto: University of Toronto Press), 1981, pp. 1–290.

Moore, G. E. 1903. *Principia ethica*, (Cambridge, Cambridge University Press).

Norcross, A. 1997. 'Good and bad actions', *Philosophical Review* 106:1–34.

Parfit, D. 1984. *Reasons and persons*, (Oxford: Clarendon Press).

Pettit, P. and Brennan, G. 1986. 'Restrictive consequentialism', *Australasian Journal of Philosophy* 64:438–55.

Pettit, P. and Smith, M. 2000. 'Global consequentialism' in Brad Hooker, Elinor Mason and Dale E. Miller (eds.), *Morality, rules and consequences: a critical reader*, (Edinburgh: Edinburgh University Press), pp. 121–133.

Railton, P. 1984. 'Alienation, consequentialism, and the demands of morality', *Philosophy and Public Affairs* 13:134–71.

Railton, P. 1988. 'How thinking about character and utilitarianism might lead to rethinking the character of utilitarianism', *Midwest Studies In Philosophy* 13:398–416.

Rawls, J. 1971. *A theory of justice*, (Cambridge Mass.: Harvard University Press).

Sidgwick, H. 1907. *The methods of ethics*, 7th edition, (London: Macmillan).

Smart, J. J. C. 1956. 'Extreme and restricted utilitarianism', *Philosophical Quarterly* 6:344–54.

Smart, J. J. C. 1973. 'Outline of a system of utilitarian ethics', in Smart and Williams, *Utilitarianism: for and against*, pp. 3–74.

Smart, J. J. C. and Williams, B. A. O. 1973. *Utilitarianism: for and against*, (Cambridge: Cambridge University Press).

Smith, H. 1988. 'Making moral decisions', *Noûs* 22:89–108.

Smith, H. 1991. 'Deciding how to decide: is there a regress problem?' in Michael Bacharach and Susan Hurley (eds.), *Foundations of decision theory*, (Oxford: Basil Blackwell), pp. 194–219.

Stocker, M. 1976. 'The schizophrenia of modern ethical theories', *Journal of Philosophy* 73:453–466.

Streumer, B. 2003. 'Can consequentialism cover everything?', *Utilitas* 15:237–47.

Streumer, B. 2005. 'Semi-global consequentialism and blameless wrongdoing: reply to Brown', *Utilitas* 17:226–230.

Williams, B. A. O. 1973. 'A critique of utilitarianism', in Smart and Williams, *Utilitarianism: for and against*, pp. 77–150.

Zimmerman, M. 1996. The concept of moral obligation, (Cambridge: Cambridge University Press).