

AS



Satisficing Consequentialism

Author(s): Michael Slote and Philip Pettit

Source: *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol. 58 (1984), pp. 139-163+165-176

Published by: Blackwell Publishing on behalf of The Aristotelian Society

Stable URL: <http://www.jstor.org/stable/4106846>

Accessed: 15/10/2008 09:26

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



The Aristotelian Society and Blackwell Publishing are collaborating with JSTOR to digitize, preserve and extend access to Proceedings of the Aristotelian Society, Supplementary Volumes.

<http://www.jstor.org>

SATISFICING CONSEQUENTIALISM

Michael Slote and Philip Pettit

I—Michael Slote

Act-consequentialism is generally characterized as a certain sort of view about the relation between an act's rightness and its consequences. An act-consequentialist holds that states of affairs (outcomes, consequences) can be objectively or impersonally ranked according to their goodness and that any given act is morally right or permissible if and only if its consequences are at least as good, according to the impersonal ranking, as those of any alternative act open to the agent—the doing of an act being itself included among its consequences.¹ An act-utilitarian is, according to the prevalent conception, an act-consequentialist with a particular view about how states of affairs are to be impersonally ranked: roughly speaking, the goodness of states of affairs depends only on the well-being, happiness, satisfaction, utility, or desire-fulfillment of the individuals who exist in those states of affairs and one state of affairs is better than another just in case it contains a greater sum of individual utilities, or a greater overall balance of satisfaction over dissatisfaction.² Thus act-consequentialism holds that a right act must be optimistic, and the act-utilitarian, in addition, that optimizing always means *maximizing* the sum of individual well-being, desire-fulfillment, etc. But these theses need not go together. Nowadays, it is by no means unusual for an (act-)consequentialist not to be a utilitarian and to hold, for example, that considerations of justice may affect the goodness of overall states of affairs without affecting the sum total of individual utilities.

Act-consequentialism itself, on the other hand, has been seen as a unitary moral conception by both defenders and critics. But

¹ For present purposes it will not, I think, be important to distinguish between acts and courses of action.

² This characterization ignores various forms of average utilitarianism. For a technically more precise elaboration of some of these distinctions, see A. K. Sen, 'Utilitarianism and Welfarism', *Journal of Philosophy* 76, 1979, pp. 463ff.

the claim that the rightness of an act depends on whether it produces the best consequences impersonally judged can in fact be broken down into a pair of claims that need not go together, and the major purpose of this essay will be to show how this is possible and, in consequence, to suggest a useful widening of the notion of (act-)consequentialism. The idea that the rightness of an act depends solely on its consequences, i.e., on how (impersonally) good its consequences are, is separable from the idea that the rightness of an act depends on its having the best consequences (producible in the circumstances); the second thesis entails the first, but not *vice versa*, yet standard conceptions of consequentialism entail both these theses. Roughly, then, consequentialism standardly involves the claim that the rightness of acts depends on whether their consequences are good enough together with the particular view that only the best possible (in certain circumstances) is good enough. And given this way of partitioning standard consequentialism, it is not perhaps immediately obvious why these two theses should naturally or inevitably go together. Could not someone who held that rightness depended solely on how good an act's consequences were also want to hold that less than the best was sometimes good enough, hold, in other words, that an act might qualify as morally right through having good enough consequences, even though better consequences could have been produced in the circumstances?

In what follows I shall try to give some of the reasons why someone might want to hold just this sort of view. It is a view that, to the best of my knowledge, has not been explicitly suggested previously; but I hope to show not only that there is nothing incoherent about it but also that it has attractive features lacking in standard act-consequentialism. Furthermore, it seems terminologically natural to treat any view that makes rightness depend solely on the goodness of consequences as a form of consequentialism, so once the feasibility of the idea that less than the best may be good enough becomes apparent, it will be appropriate to treat the view that rightness depends on whether consequences are good enough and that less than the best may sometimes be good enough as a form of consequentialism. The traditional or standard view that rightness depends on whether the consequences are the best producible in the

circumstances will then most naturally be seen as a particular kind of consequentialism, rather than as constituting consequentialism *per se*.³ And it will be natural to characterize this particular kind of consequentialism as 'optimizing consequentialism' since it holds that rightness depends on whether consequences are good enough and that only the best is good enough. By contrast, the new sort of consequentialist view just mentioned might appropriately be labelled 'satisficing consequentialism', if we may borrow from the recent literature of economics, where the notion of satisficing has been used to express the idea that (rational) economic agents may sometimes choose what is good enough, without regard for whether what they have chosen is the best thing (outcome) available in the circumstances.

Now the idea of satisficing (utilitarian) consequentialism deserves to be explored as a formal possibility quite apart from its intuitiveness or ultimate supportability; but in fact it can be made to appear of more than formal interest. Even those opposed to consequentialism and utilitarianism as moral theories have tended to think that (extramoral individualistic) rationality requires an individual to maximize his satisfactions or do what is best for himself;⁴ but the recent economics literature concerning satisficing suggests the possibility of a non-optimizing form of individual rational choice, and by giving a brief philosophical elaboration of the idea of rational satisficing, I hope to make the idea of *moral* satisficing, and satisficing

³ Surely, it would be terminologically odd to treat someone who affirmed what I am going to call 'satisficing consequentialism' as *denying* consequentialism, yet this is precisely what present-day terminology requires—see, for example, Bernard Williams 'A Critique of Utilitarianism', in Smart and Williams, *Utilitarianism: For and Against*, Cambridge University Press, 1973, p. 90.

Incidentally, the possibility of satisficing (act-)consequentialism is (I believe) unintentionally suggested by things said in the Introduction to Sen and Williams, eds., *Utilitarianism and Beyond* (Cambridge University Press, 1982, p. 3f.). At one point, consequentialism is simply characterized as a theory 'which claims that actions are to be chosen on the basis of the states of affairs which are their consequences . . .' and this is neutral as between optimizing and satisficing forms of consequentialism. However, the Introduction goes on to treat consequentialism as involving the production of optimal consequences and the possibility of consequentialist moral satisficing is never mentioned. Certainly, in Sen, *op. cit.*, consequentialism is *defined* in terms of optimal consequences.

⁴ See, e.g., Rawls *A Theory of Justice*, Cambridge: Harvard, 1971, pp. 23ff., 416ff.

consequentialism in particular, seem more attractive.⁵ It will turn out, furthermore, that ordinary or common-sense morality also regards acts that are less than the best (most beneficent) possible as sometimes good enough and so not morally wrong even apart from any sacrifices a better (more beneficent) act might require from the agent. And to the degree that common-sense morality allows for ‘moral satisficing’ in the area of beneficence (benevolence), the possibility of a satisficing form of (utilitarian) consequentialism is also underscored and made more appealing. After all, consequentialists have long sought for ways—rule utilitarianism, probabilistic act-utilitarianism, etc.—of reconciling their views (making them seem less out of line) with common-sense morality, and we shall see towards the end of this essay that satisficing consequentialism has a number of advantages, in terms of common-sense moral plausibility, over optimizing forms of consequentialism, utilitarian and non-utilitarian alike. We shall also see that prominent views about the object of and motivation behind morality that are taken to support optimizing consequentialism or utilitarianism support these views only in their most general form and are equally consistent with optimizing or satisficing versions. But first to the idea of rational individual satisficing.

I

Consider an example borrowed from the satisficing literature of economics, but treated in such a way as to emphasize its relevance to philosophical discussions of rationality, rather than its implications for economic theory. An individual planning to move to a new location and having to sell his house may seek, not to maximize his profit on the house, not to get the best price for it he is likely to receive within some appropriate time period, but simply to obtain what he takes to be a good or satisfactory price. What he deems satisfactory may depend, among other things,

⁵ For relevant discussions in the economics literature of satisficing, see, e.g., H. Simon, ‘A Behavioral Model of Rational Choice’, *Quarterly Journal of Economics* 69, 1955, pp. 99–118; Simon, ‘Theories of Decision Making in Economics and Behavioral Science’, *American Economic Review* XLIX, 1959, pp. 253–83, *Administrative Behavior*, N.Y.: Macmillan, 1961, second edition; and R. Ewert and J. March, eds., *A Behavioral Theory of the Firm*, Englewood Cliffs, N.J.: Prentice-Hall, 1963.

on what he paid for the house, what houses cost in the place where he is relocating, and on what houses like his normally sell at. But given some notion of what would be a good or satisfactory price to sell at, he may fix the price of his house at that point, rather than attempting, by setting it somewhat higher, to do better than that, or do the best he can. His reason for not setting the price higher will not, in that case, be some sort of anxiety about not being able to sell the house at all or some feeling that trying to do better would likely not be worth the effort of figuring out how to get a better price. Nor is he so rich that any extra money he received for the house would be practically meaningless in terms of marginal utility. Rather he is a 'satisficer' content with good enough and does not seek to maximize (optimize) his expectations. His desires, his needs, are moderate, and perhaps knowing this about himself, he may not be particularly interested in doing better for himself than he is likely to do by selling at a merely satisfactory price. If someone pointed out that it would be better for him to get more money, he would reply, not by disagreeing, but by pointing out that for him at least a good enough price is good enough.

Such a person clearly fails to exemplify the maximizing and optimizing model of individual rationality advocated by utilitarians like Sidgwick and anti-utilitarians like Rawls. But I think he nonetheless represents a possible idea of (one kind of) individual rationality, and the literature of economic satisficing in the main treats such examples, both as regards individuals and as regards economic units like the firm, as exemplifying a form of rational behavior. It might be possible to hold on to an optimizing or maximizing model of rationality and regard satisficing examples as indications of the enormous prevalence of irrational human behavior, but this has typically not been done by economists, and I think philosophers would have even less reason to do so. For there are many other cases where satisficing seems rational, or at least not irrational, and although some of these are purely hypothetical, hypothetical examples are the stock-in-trade of ethical and moral-psychological theory even when they are of little or no interest to economists.

Imagine that it is mid-afternoon; you had a good lunch, and you are not now hungry; neither, on the other hand, are you sated. You would enjoy a candy bar or Coca Cola, if you had

one, and there is in fact, right next to your desk, a refrigerator stocked with such snacks and provided gratis by the company for which you work. Realizing all this, do you, then, necessarily take and consume a snack? If you do not, is that necessarily because you are afraid to spoil your dinner, because you are on a diet or because you are too busy? I think not. You may simply not feel the need for any such snack. You turn down a good thing, a sure satisfaction, because you are perfectly satisfied as you are. Most of us are often in situations of this sort, and many of us would often do the same thing. We are not boundless optimizers or maximizers, but are sometimes (more) modest in our desires and needs. But such modesty, such moderation, need not be irrational or unreasonable on our part.

In the example just mentioned, moderation is not functioning as a means to greater overall satisfactions and is thus quite different from the instrumental virtue recommended by the Epicureans. The sort of moderation I am talking about is not for the sake of anything else, indeed it may not be for its own sake either, if by that is meant that it is some sort of admirable trait or virtue. If one has the habit of not trying to eke out the last possible satisfaction from situations and of resting content with some reasonable quantity that is less than the most or best one can do, then one has a habit of moderation or modesty as regards one's desires and satisfactions, and it may not be irrational to have such a habit, even if (one recognizes that) the contrary habit of maximizing may also not be irrational. But if a maximizer (optimizer) lacking the habit of moderation in the above sense need not be immoderate in that ordinary sense of the term that implies unreasonableness, then the habit of being satisfied with less than the most or best may not be a virtue, even if such moderation is also neither irrational nor an anti-virtue.⁶

⁶ A person whose desires are moderate or modest might also be called 'temperate', but what we have been saying about the habit of moderation is very different from what Aristotle says about what he calls the virtue of temperance. On Aristotle's conception a temperate individual has the right amount of desire for the right sorts of things, etc., and such rightness, roughly, involves a mean between two less right extremes. But our talk of moderation in the text above is not supposed to imply that taking more than moderation would allow is in any way wrong or unreasonable. Moderation involves a mean between extremes that are not necessarily (both) more undesirable than moderation itself is. And it is possible to use the notion of 'temperance' or 'temperateness' in a similarly non-Aristotelian and (relatively) value-free way.

But if there is nothing irrational or unreasonable about maximizing, isn't the moderate individual who is content with less a kind of ascetic? Not necessarily. An ascetic is someone who, within certain limits, minimizes his enjoyments or satisfactions; he deliberately leaves himself with less, unsatisfied. The moderate individual, on the other hand, is someone content with (what he considers) a reasonable amount of satisfaction; he wants to be satisfied and up to a certain point he wants more satisfactions rather than fewer, to be better off rather than worse off; but there is a point beyond which he has no desire, and even refuses, to go. There is a space between asceticism and the attempt to maximize satisfactions, do the best one can for oneself, a space occupied by the habit (if not the virtue) of moderation. And because such moderation is not a form of asceticism, it is difficult to see why it should count as irrational from the standpoint of egoistic or extra-moral individual rationality.⁷

Now the kind of example just mentioned differs from the case of satisficing house selling in being independent of any monetary transaction. But the example differs importantly in another way from examples of satisficing mentioned in the literature of economics. Economists who have advocated the model of rational satisficing for individuals, firms, or state bodies have pointed out that, quite independently of the costs of gaining further information or effecting new policies, an entrepreneur or firm may simply seek a satisfactory return on investment, a satisfactory share of the market, a satisfactory level of sales, rather than attempting to maximize or optimize under any of these headings. But this idea of rational satisficing implies only that individuals or firms *do not* always *seek* to optimize and are *satisfied* with attaining a certain 'aspiration level' less than the best that might be envisaged. It does not imply that it could be rational actually to *reject* the better for the good enough in situations where both were available. In the example of house selling, the individual accepts less than he might well be able to get, but he doesn't accept a lower price when a higher bidder makes an equally firm offer. And writers on satisficing generally

⁷ *Rational* satisficing seems to involve not only a disinclination to optimize, but a reasonable sense of when one has enough. To be content with much less than one should be is (can be) one form of *bathos*.

seem to hold that satisficing only makes sense as a habit of not seeking what is better or best, rather than as a habit of actually rejecting the better, when it is clearly available, for the good enough. Thus Herbert Simon, in 'Theories of Decision Making . . .' (*loc. cit.* n. 5), develops the idea of aspiration level and of satisficing, but goes on to say that 'when a firm has alternatives open to it that are at or above its aspiration level, it will choose the best of those known to be available'.

However, the example of the afternoon snack challenges the idea that the satisficing individual will never explicitly reject the better for the good enough. For the individual in question turns down an immediately available satisfaction, something he knows he will enjoy. He isn't merely not trying for a maximum of satisfactions, but is explicitly rejecting such a maximum. (It may be easier to see the explicitness of the rejection if we change the example so that he is actually offered a snack by someone and replies: no thank you, I'm just fine as I am.) And I think that most of us would argue that there is nothing irrational here. Many of us, most of us, occasionally reject afternoon snacks, second cups of tea, etc., not out of (unconscious) asceticism, but because (to some degree) we have a habit of moderation with regard to certain satisfactions. The hypothetical example of the afternoon snack thus takes the idea of rational satisficing a step beyond where economists, to the best of my knowledge, have been willing to go.

At this point, however, it may be objected that the example may be one of rational behavior but is less than clear as an example of satisficing. The individual in question prefers not to have a certain satisfaction and certainly deliberately rejects the maximization of satisfactions, if we think of satisfactions as like pleasures or enjoyments. But to the extent the individual rejects an available satisfaction, he presumably shows himself to prefer (or desire) not to have that satisfaction and so in some (trivial?) sense is maximizing the satisfaction of his preferences (or desires). More importantly, perhaps, it is not clear that the moderate individual must think of himself as missing out on anything *good* when he forgoes the afternoon snack. For although he knows he would enjoy the snack, the very fact that he rejects such enjoyment might easily be taken as evidence that he doesn't in the circumstances regard such enjoyment as a good

thing. In that case, he may be satisficing in terms of some quantitative notion of satisfaction, but not with respect to some more refined or flexible notion of (his own) individual good, and the example would only provide a counter-example to a rather crude maximizing ideal of rationality, not to the idea that it is irrational to choose what is less good for one when something better is available.

However, even if the enjoyment of a snack does count as a rejected personal good for the individual of our example, that fact may be obscured, both for him and for us, by the very smallness or triviality of the good in question. And so in order to deal with our doubts, it may, then, be useful at this point to consider other examples, more purely hypothetical than the present one, where the good forgone through satisficing is fairly obvious.

How do we react to fairy tales in which the hero or heroine, offered a single wish, asks for a pot of gold, for a million (1900) dollars, or, simply, for (enough money to enable) his family and himself to be comfortably well off for the rest of their lives. In each case the person asks for less than he might have asked for, but we are not typically struck by the thought that he was *irrational* to ask for less than he could have, and neither, in general, do the fairy tales themselves imply a criticism of this sort; so, given the tendency of such tales to be full of moralism about human folly, we have, I think, some evidence that such fairy-tale wishes need not be regarded as irrational. (In not regarding them as irrational, we need not be confusing what we know *about* fairy-tale wishes with what the individual *in* a given fairy tale ought to know.)

Now the individual in the fairy tale who wishes for *less* than he could presumably exemplifies the sort of moderation discussed earlier. He may think that a pot of gold or enough money to live comfortably is all he needs to be satisfied, that anything more is of no particular importance to him. At the same time, however, he may realize (be willing to admit) that he could do better for himself by asking for more. He needn't imagine himself constitutionally incapable of benefitting from additional money or gold, for the idea that one will be happy, or satisfied, with a certain level of existence by no means precludes the thought (though it perhaps precludes *dwelling* on the thought) that one will not be as well off as one could be. It merely precludes the

sense of wanting or needing more for oneself. Indeed the very fact that someone could actually explicitly wish for enough money to be comfortably well-off is itself sufficient evidence of what I am saying. Someone who makes such a wish clearly acknowledges the possibility of being better off and yet chooses—knowingly and in some sense deliberately chooses—a lesser but personally satisfying degree of well-being.⁸ And it is precisely because the stakes are so large in such cases of wishing that they provide clearcut examples of presumably rational individual satisficing. But, again, the sort of satisficing involved is not (merely) the kind familiar in the economics literature where an individual seeks something other than optimum results, but a kind of satisficing that actually rejects the available better for the available good enough. Although the individual with the wish would be better off if he wished for more, he asks for less (we may suppose that if the wish grantor prods him by asking ‘Are you sure you wouldn’t like more money than that?’, he sticks with his original request). And if we have any sympathy with the idea of moderation, of modesty, in one’s desires, we shall have to grant that the satisficing individual who wishes, e.g., for less money is not irrational. Perhaps we ourselves would not be so easily satisfied in his circumstances, but that needn’t make us think him irrational for being moderate in a way, or to a degree, that we are not.⁹

II

Given the above discussion of the nature and justification of rational satisficing, the way may be prepared for an examination

⁸ He may feel it is better that he choose less, and this entails the thought that it is (in one sense) better for him to choose less. But none of this need entail the thought that he will be personally better off if he chooses less, that such a choice will be better for him in the other natural sense of that expression. On this see my *Goods and Virtues*, Oxford: 1983, ch. 3.

Incidentally, our example requires us to assume that the wisher’s choice is not influenced by a (reasonable) fear of being corrupted by getting more than he in fact asks for; but his non-optimizing attitude in fact shows a *certain kind* of present non-corruption.

⁹ In fact, it is hard to see how any specific monetary wish can be optimizing if the individual is unsure about his own marginal utility curve for the use of money. To that extent, we are all *necessarily* satisficers in situations where we can wish for whatever we want, unless, perhaps, we are allowed to wish for our own greatest future well-being in those very terms. If satisficing were irrational, would that mean that anything other than such an explicitly optimizing wish would be irrational?

of moral satisficing. But I shall not immediately proceed to a discussion of (the varieties of) satisficing consequentialism, because I believe we can make the strongest case for this new form of consequentialism by first pointing out the non-optimizing character of the common-sense morality of benevolence (beneficence).

Consider a manager of a resort hotel who discovers, late one evening, that a car has broken down right outside its premises. In the car are a poor family of four who haven't the money to rent a cabin or buy a meal at the hotel, but the manager offers them a cabin gratis, assuming (as we may assume for the sake of argument) that it would be wrong not to do so. In acting thus benevolently, however, she doesn't go through the complete list of all the empty cabins in order to put them in the best cabin available. She simply goes through the list of cabins till she finds a cabin in good repair that is large enough to suit the family. Imagine, further, that, as with examples of rational satisficing from the economics literature, she chooses the cabin she does because it seems a satisfactory choice, good enough, not because, as an optimizer, she thinks that further search through the list of cabins will not be worth it in terms of time expended and the likelihood of finding a (sufficiently) better cabin. In such circumstances, optimizing act-consequentialism or act-utilitarianism would presumably hold that the manager should look further for a better room. (Assume there is a better room and that she will easily find it if she proceeds further through the list.) But I think ordinary morality would regard her actions as benevolent and her choice of a particular room for the family in question as morally acceptable, not wrong. She may not display the optimizing benevolence that standard act-consequentialism would require, under the circumstances, but in ordinary moral terms she has done well enough by the family that is stranded and had no obligation to do any better.

The example illustrates the possibility of a morally acceptable satisficing benevolence that does not seek to optimize with respect to those benefitted (or those affected) by one's actions. But our earlier examples of rational individual satisficing extended the notion beyond the usual examples from the literature of economics to include cases where someone explicitly rejects a better (or the best) alternative; and the same possibility

in fact also exists in the area of moral satisficing. Thus consider again our hotel manager and the travellers she benefits. They have now moved to the cabin she has found for them and are all hungry. But it is late; so the manager tells the lone remaining waiter in the restaurant to bring out a meal for the travellers from among the dishes that remain from dinner and that will not be usable the next day. Assume that there are a large variety of dishes, some more luxurious or splendid than others, and that the waiter asks what, among these things, he should bring the newly arrived travellers. The hotel manager may say: oh, just something good and substantial, it needn't be too fancy or elaborate. Alternatively, the waiter may ask whether he should bring them the 'special dinner' and the manager may say: no, there's no need for anything that fancy, just bring them something appetizing and good. In either case, the manager seems deliberately to be rejecting an alternative that stands a good chance of being preferable to the poor family in question. Most people prefer the 'special dinner', apart from its price, and the manager has no reason to believe that her new guests are particularly moderate or modest in their desires. Yet her reason for choosing as she does may not be consideration for the waiter, who may have the same amount of work to do whatever he brings the family to eat, nor even less a snobbish sense of charity that regards the 'special dinner' as too good for the family in question. Rather, she may be expressing in her benevolent actions a kind of moderation that she may also evince in her self-regarding choices. And, again, I think common-sense would regard such deliberately non-optimific benevolence as morally acceptable, not wrong.

In addition, if I may appeal again to an even more hypothetical example in order to underscore the similarity with what was said earlier about self-regarding individual rationality, consider a fairy-tale wish regarding people other than oneself. A warrior has fought meritoriously and died in a good cause, and the gods wish to grant him a single wish for those he leaves behind, before he enters Paradise and ceases to be concerned with his previous life. Presented with such an opportunity, may not the warrior wish for *his family* to be comfortably well off forever after? And will we from a common-sense standpoint consider him to have acted wrongly or non-benevolently

towards his family because he (presumably knowingly) rejected an expectably better lot for them in favor of what was simply good enough? Surely not.

But the warrior and hotel manager examples not only offer further illustration of the idea of satisficing benevolence, but also help to make clear that common-sense morality differs from standard optimizing consequentialism with regard to the morality of benevolence quite apart from issues concerning the amount of sacrifice one may correctly require from moral agents. In many familiar cases where (optimizing) act-consequentialism and act-utilitarianism diverge from common-sense morality with regard to what an agent may permissibly do, the former require that an agent benefit others even though doing so requires him to make a large personal sacrifice, whereas common-sense morality regards the agent as permitted to refrain from making such a sacrifice. But in the fairy-tale example of the dying warrior, the warrior who chooses less than the best for his family does not do so because a choice of something better would require too great a sacrifice. Neither choice would require *any* sort of personal sacrifice. And by the same token the hotel manager's personal sacrifice (if any) presumably stays constant however splendid a meal she decides to give the poor travellers. So the divergence between common-sense morality and standard (utilitarian) act-consequentialism with regard to such cases cannot be accounted for in terms of a disagreement over whether one can correctly require an agent to sacrifice his own desires, projects and concerns in the name of overall optimality. With regard to such cases they disagree, rather, as to how much good an agent may be morally required to do (for others) given a total absence, or constant amount, of agent sacrifice. Optimizing (utilitarian) act-consequentialism will hold that the moral agent must produce the best possible results in such circumstances, ordinary morality that producing sufficiently good non-optimal results may be all that is required.

Of course, this is not the only way in which common-sense morality diverges from optimizing act-consequentialism: ordinary morality also contains deontological restrictions on what an agent may permissibly do in the name of overall optimality. But the fact that common-sense morality allows a satisficing concern for good results, a less than optimific

beneficence, to be permissible in some cases where deontological restrictions are irrelevant (and where there is no issue of personal sacrifice on the part of the agent) suggests the possibility of a satisficing form of pure act-consequentialism. And since the plausibility of various forms of consequentialism partly depends on how far their implications diverge from the deliverances of ordinary moral intuition, this new form of consequentialism may turn out to have some distinctive advantages over traditional optimizing forms of consequentialism.

The idea of satisficing act-consequentialism is not, in fact, entirely new. It is to a certain extent anticipated, for example, by the sort of 'negative utilitarianism', proposed briefly by Karl Popper in *The Open Society and its Enemies*, according to which we have a moral duty to minimize suffering and evil, but no general duty to maximize human happiness. In the course of defending this doctrine, Popper claims, in particular, that the idea of relieving suffering has a greater moral appeal to us than the idea of increasing the happiness of a man who is doing well already. So Popper not only offers an example of non-optimizing consequentialism, but indicates that it can be based on 'satisficing' common-sense moral intuitions of the kind we ourselves mentioned above.¹⁰

But Popper suggests only one form of non-optimizing consequentialism, and it is a form attended by a number of serious difficulties. The presumed fact that adding to happiness has less moral appeal than relieving suffering hardly implies, for example, that our only duty is to relieve suffering. And the latter idea, which constitutes the essence of negative utilitarianism, seems to have the absurd consequence that we would do all that duty requires, if we painlessly destroyed all of suffering humanity.¹¹ Negative utilitarianism seems, then, to entail unacceptable views about when less than the best possible is good enough. But by virtue of its asymmetric treatment of

¹⁰ See *The Open Society and its Enemies*, London: Routledge and Kegan Paul, 1974, vol. I, ch. 5, note 6; chapter 9, note 2. For another discussion of our common-sense satisficing moral intuitions, see R. M. Adams, 'Must God Create the Best?', *Philosophical Review* 81, 1972, pp. 317-32.

¹¹ The point is made by R. N. Smart in 'Negative Utilitarianism', *Mind* 67, 1958, p. 542f.

human happiness and suffering it also rules out the possibility of morally permissible satisficing throughout a wide range of cases. Wherever the relief of suffering is in question, it demands that suffering be minimized; but common-sense morality is not in fact so demanding in this respect. A medic attending the wounded on the battlefield may attend to the first (sufficiently) badly wounded person he sees without considering whether there may be someone in even worse shape nearby, and from a common-sense moral standpoint such behavior seems perfectly acceptable. So although Popper's variety of satisficing consequentialism involves a fundamental asymmetry between good and evil, happiness and suffering, a less asymmetric form of such consequentialism may, for a number of reasons, be more attractive and intuitive. What was appealing, in our discussion of rational individual satisficing, was the idea of a reasonable sufficiency of good less than the best attainable. But this idea gains only imperfect expression in a moral theory like negative utilitarianism, which, among other things, treats the eradication of our suffering race as 'good enough' but not the behavior of the morally satisficing medic. What may be needed is a form of satisficing consequentialism with a more plausible conception of what counts as good enough, and it will help us towards such a theory if we consider another form of satisficing consequentialism that can be found in one of the great classics of utilitarian moral philosophy, Bentham's *An Introduction to the Principles of Morals and Legislation*.

The 1823 edition of the *Introduction* differs from that published in 1789 principally in regard to some clarificatory notes, and one of these notes discusses the Principle of Utility and interprets it as requiring that everyone seek the greatest happiness of those affected by his actions. But the unaltered main text of the book treats the Principle of Utility in quite a different (indeed incompatible) way as 'that principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question: or, what is the same thing in other words, to promote or to oppose that happiness'.¹² The discussion in footnote presents a typical

¹² See the J. H. Burns and H. L. A. Hart edition of the *Introduction*, London: Methuen, 1970, pp. 11f. However, also see p. 282.

optimizing form of (utilitarian) act-consequentialism; but the earlier main text says nothing about 'best' or 'greatest good', and since an act may promote happiness without producing the most happiness possible, in given circumstances, the earlier text presents a kind of satisficing utilitarian act-consequentialism. [As with Popper's discussion, a particular form of satisficing consequentialism is advocated without the *general idea* of satisficing (act-)consequentialism or (act-)utilitarianism being mentioned.]

Now the form of satisficing consequentialism Bentham advocates has some unfortunate features. If we take it quite literally, his theory treats the rightness of any action as dependent solely on the results of that single action; an act is right even if it adds very little to the sum of human happiness (the net balance of happiness over unhappiness) and even if an alternative is available which is much more productive of happiness. By the same token, an act is wrong if it subtracts from the sum of human happiness, even when every available alternative has worse results. (Given utilitarian notions of consequences, it is easy to imagine such cases.) These implications do not square with ordinary moral intuitions, and they are due to the *non-comparativeness* of the just-mentioned form of consequentialism. Bentham's satisficing consequentialism regards an act as having produced *enough* good or happiness to be right if it favourably alters the balance of happiness over unhappiness even to the *slightest* extent, but if an alternative is available which would produce much more good, we should perhaps not normally feel that a slight addition to happiness was (morally) good enough. The sufficiency of a slight contribution to happiness would somewhat depend on what else was available, and a comparative form of satisficing consequentialism would require such alternatives to be taken into account in judging what was good enough. An act producing only slight good might not be judged good enough, even if some non-optimific alternative act producing a great deal of good were regarded as such. (The medic of our example would presumably be wrong to go around simply applying bandages, even if that would slightly ameliorate the situation; but he might still be morally permitted to satisfice in the way mentioned earlier.)

Similarly, where an agent cannot avoid acting in such a way

as will lead to the decrease of human happiness, a non-comparative form of satisficing consequentialism, like Bentham's, must presumably treat whatever the agent does as wrong.¹³ But it seems more plausible to take into consideration the alternatives to a given act and regard an act as having (circumstantially) good enough consequences if all its alternatives have worse consequences (for the sum of human happiness). So (some of) the difficulties of Bentham's satisficing consequentialist theory of right action are due to its insensitivity to the consequences of alternative actions, and we are thus pointed in the direction of some form of *comparative* satisficing act-consequentialism, in our search for a viable alternative to standard optimizing act-consequentialism.

Given, furthermore, our earlier plausible examples of moral satisficing with respect to the relief of suffering, it would seem that satisficing act-consequentialism does best to avoid the radical asymmetry between suffering and happiness entailed by Popper's negative utilitarianism, while at the same time remaining sensitive to the consequences of alternative actions. But having said as much, it would perhaps at this point be a good idea to consider some of the relative advantages and disadvantages of (plausible versions of) satisficing and optimizing act-consequentialism.

III

Some of the relative strengths of satisficing consequentialism most clearly appear in certain kinds of cases where an individual can through his own efforts do a great deal to relieve great human suffering. Consider a doctor who wants to help mankind, but is for personal reasons particularly affected by the plight of people in India—perhaps he is attracted to Indian art or religion or is very knowledgeable about the history of India. Now an optimizing (utilitarian) act-consequentialist would presumably say that a doctor who volunteered to work in India should

¹³ Recent moral philosophy has taken very seriously the possibility that an agent might through no fault of his own be put in a situation of moral tragedy where he cannot avoid acting wrongly. But surely there are cases where anything one does will have bad consequences (in the permissive utilitarian sense of the term), yet where it is not wrong to perform an act with less bad consequences than any feasible alternative.

consider whether the suffering there is worse than in other countries (and the opportunities to help great enough) so that by going to India he is likely to do more good than he can do elsewhere. And aside from the doctor's motivation and thought processes, it will simply be wrong of him to go to India if some other course of action would do more good for mankind, on any usual optimizing act-consequentialist conception. But many people who wish to relieve human suffering do not consider whether their actions are likely to produce the greatest amount of good possible. When they find a course of action that they think will make a great (enough) contribution to the relief of suffering and that can compel their personal allegiance and energies, they may act accordingly, without considering whether they might not do more good elsewhere. And such moral satisficing does not, from a common-sense moral standpoint, seem wrong; a person who has done a great deal to relieve suffering would not normally be thought to have acted wrongly because she could have done even greater service elsewhere.

Now the reason why, in the present sort of case, less than the best can seem good enough may have something to do with how much good the satisficing individual does and/or aims to do. The good he does is sufficiently great and sufficiently close to the best he could do so as to make it implausible to deny the rightness of his action(s). So for such situations it may well be possible formally to elaborate the notion of enoughness as some sort of percentage or other mathematical function of the best results attainable by the agent. I shall not attempt to spell out the details of any particular plausible way in which this might be attempted. But if satisficing consequentialism has the sort of initial plausibility I think it has, then the way will be open to such a formal elaboration of enoughness and to a consequent precision (over large ranges of cases) about what counts as a morally permissible level of act-consequentialist satisficing.

According to such satisficing act-consequentialism, then, an agent may permissibly choose a course of action that seems to him to do a great deal or a sufficient amount towards the relief of human suffering without considering whether such action is optimific in the relief of suffering among all the possible courses of action open to him and without the course of action he chooses actually being optimific. His sense of what counts as doing a

great deal, or a sufficient amount, towards the relief of suffering may in part reflect what he knows about the most good that agents can do in circumstances like his—and what he knows generally about the world—but in doing the great good he does he need not seek to do or actually succeed in doing the most good possible. And, as we have seen, such a view of what an agent may permissibly do comes closer to ordinary moral views about benevolence than the usual forms of optimizing act-consequentialism.

One of the chief implausibilities of traditional (utilitarian) act-consequentialism has been its inability to accommodate moral supererogation. But a satisficing theory that allows less than the best to be morally permissible can treat it as supererogatory (and especially praiseworthy) for an agent to do more good than would be sufficient to insure the rightness of his actions. Thus, if the person with special interest in India sacrifices that interest in order to go somewhere else where he can do even more good, then he does better than (some plausible version of) satisficing act-consequentialism requires and acts supererogatorily. But optimizing act-consequentialism will presumably not treat such action as supererogatory because of its (from a common-sense standpoint) inordinately strict requirements of benevolence.

Moreover, critics of optimizing consequentialism have recently tended to focus on one particular way in which such consequentialism implausibly offends against common-sense views of our obligations of beneficence. They have pointed out that (optimizing) act-consequentialism makes excessive demands on the moral individual by requiring that she abandon her deepest commitments and projects whenever these do not serve overall impersonally judged optimality. For example, it has been held by Samuel Scheffler (and others) that it is unfair or unreasonable to demand such sacrifice of moral agents, and by Bernard Williams (and others) that such requirements alienate individuals from their own deepest identities as given in the projects and commitments they hold most dear, thus constituting attacks on their integrity (integralness) as persons.¹⁴ In the wake

¹⁴ See, for example, Scheffler's *The Rejection of Consequentialism*, Oxford, 1982; and Williams in Smart and Williams, *op. cit.*

of these criticisms, many philosophers have advocated rejecting act-consequentialism in favor of some more commonsensical moral view that makes it permissible to pursue non-optimific personal projects and commitments.

What I would like to suggest here, however, is that at least some of this moral accommodation of individual desires, commitments and projects can be accomplished within a (utilitarian) act-consequentialist framework. Satisficing act-consequentialism can permit a doctor to work in India, even if he could do more good elsewhere, as long as the amount of good he will do in India is judged to be sufficient. And this then permits the doctor to satisfy his special interest in or concern for things Indian while at the same time fulfilling all that morality demands of him. Similarly, a person interested in pure laboratory research might be permitted to pursue such research if it were likely to yield great practical benefits for mankind (rather than threaten human survival). Such consequentialism in effect then allows various sorts of *compromise* between the demands of impersonal morality and personal desires and commitments. To that extent, it allows greater scope for personal preferences and projects than traditional optimizing act-consequentialism does. However, it offers less scope than would be available on most common-sense views of what an agent may permissibly do. For ordinary morality would presumably allow an agent (capable of doing better) to pursue projects that do not contribute very much to overall human well-being, and satisficing consequentialism—unless it maintains a very weak view about what it is to do enough good—will rule such projects out.

However, none of the moral theories just mentioned offers an all-or-nothing solution to the problem of balancing personal projects and commitments against impersonal good. Although the point is somewhat obscured by Williams, even optimizing act-consequentialism allows for certain sorts of personal integrity, namely, integrity constituted by the desire precisely to produce the most good possible or by projects which indirectly produce the most good that a given agent can achieve. On the other hand, even anti-consequentialists who affirm some sort of common-sense permission to pursue non-optimific projects and commitments typically set limits on what sorts of projects may

be thus pursued: the desire to rise in the Mafia is presumably not among them. But within these limits satisficing act-consequentialism occupies an intermediate position; it morally accommodates more kinds of personal preference and personal integrity than traditional optimizing act-consequentialism, but fewer kinds than common-sense anti-consequentialist morality would presumably allow. As such, however, it does offer act-consequentialists the possibility of moving closer to common-sense morality and accommodates a felt need to give greater weight to personal commitments and preferences, while retaining the advantages of a (utilitarian) act-consequentialist framework.¹⁵

In addition, the choice between satisficing and optimizing consequentialism represents a genuine and difficult problem not only at the level of act evaluation, but also in regard to motives, traits of character, and everything else that can be subjected to consequentialist moral assessment. Is a good (or right) motive, for example, one that has sufficiently good (but not necessarily best) consequences among some set of relevant alternatives or is it one that has the best consequences among such alternatives? The utilitarian and consequentialist evaluation of motives goes back at least as far as Bentham's *Introduction*, where it is said that motives are good if they produce pleasure or avert pain. Clearly, this represents an embryonic form of *satisficing* motive utilitarianism.¹⁶ Something similar can also be found in Sidgwick's *The Methods of Ethics*¹⁷ and in R. M. Adams's 'Motive Utilitarianism'.¹⁸ And in point of fact satisficing forms of motive-utilitarianism and motive-consequentialism seem generally both more plausible and more interesting than any optimizing version I can think of. Act-consequentialists have

¹⁵ In 'Evaluator Relativity and Consequentialist Evaluation' (*Philosophy and Public Affairs* 12, 1983, pp. 113–32), Amartya Sen has recently suggested (roughly) that an act's rightness may depend on whether it produces consequences that are best-from-the-standpoint-of-the-agent. Such a theory, despite the lack of historic antecedents, is arguably act-consequentialist and would permit the pursuit of individual projects. Unfortunately, it also seems to make such pursuit obligatory—perhaps this could be avoided by adopting a satisficing version of the theory.

¹⁶ *Op. cit.*, p. 100.

¹⁷ Seventh edition, p. 428.

¹⁸ *Journal of Philosophy* 76, 1979, pp. 467–81.

had to grant, for example, that on many occasions a father who gives preference to his own children out of love for them may perform a morally wrong action in doing so, but they have attempted to mitigate the harshness and common-sense implausibility of that judgment by adding that the motive of paternal love may nonetheless be a morally good one because it generally leads to good consequences. The father may do what is wrong but he does so out of a morally good motive.¹⁹ However, if a motive had to be in some sense optimistic in order to count as morally good, then this irenic ascent (descent?) to the consequential morality of motives might easily fail of its purpose. Does paternal love generally produce more good than the motive of impartial benevolence, the love of mankind generally, and all other relevantly alternative single motives? If, on the other hand, we require of good motives only that they generally produce (sufficiently) good results, paternal love need not compete with these other motives in order to count as morally good, so satisficing versions of motive-consequentialism have distinct advantages, for an act-consequentialist and more generally, over optimizing versions of motive-consequentialism.

But if that is so, a question concerning the consistency of moral theories operating at different levels of evaluation immediately arises. In 'Motive Utilitarianism' Adams raises the question whether motive-utilitarianism (motive-consequentialism) is consistent with act-utilitarianism (act-consequentialism) and suggests, subject to certain qualifications, that they are. But in the context of the present discussion and relative to the assumption that one can consistently be both a motive-consequentialist and an act-consequentialist, a further question arises as to whether one can plausibly or consistently be a satisficing consequentialist with regard to some objects of consequentialist evaluation and an optimizing one with respect to others. Can one, for example, reasonably be an optimizing act-utilitarian but a satisficing motive-utilitarian?

To a greater or lesser extent, Bentham, Sidgwick, and Adams all seem to have maintained precisely this combination of views, but this may be due to their having ignored the possibility of

¹⁹ Cf. Sidgwick, *op. cit.*, p. 428; Adams, 'Motive Utilitarianism'; and R. M. Hare, *Moral Thinking*, Oxford: 1982.

satisficing act-utilitarianism rather than to any judgment that optimizing act-utilitarianism fits in well with satisficing motive-utilitarianism. (At the very end of this essay we shall briefly consider some possible reasons why satisficing consequentialism might easily be ignored or mistakenly ruled out.)

At the very least, the viability and independent plausibility of satisficing motive-consequentialism provide further motivation for satisficing act-consequentialism. They do something to allay fears that there must be something logically or conceptually wrong with any non-optimizing form of act-consequentialism and may even make it easier to regard satisficing versions of act-consequentialism as genuine competitors of the more familiar optimizing variety. Whether there may, in the end, be some sort of inconsistency or tension between optimizing act-consequentialism and satisficing motive-consequentialism is very difficult to judge, but in terms of sheer symmetry and simplicity, the superior plausibility of satisficing motive-consequentialism seems to recommend a preference for satisficing act-consequentialism as well.

IV

We come, finally, to certain views about the object (or purpose or goal) of morality and about the fundamental or ideal nature of moral motivation that may be thought to favor, indeed to mandate, optimizing, rather than satisficing, forms of (utilitarian) consequentialism. It has frequently been said, for example, that the object of morality is the general good or universal happiness, and such a view (barring any doubts one may have about morality's having any purpose or object at all) seems to imply that moral schemes—whether act-utilitarian, rule-utilitarian, motive-utilitarian, or any combination of these—must be optimific, e.g., directed towards the greatest happiness of the greatest number. (The problem whether this is best achieved by agent-indirection has no immediate relevance to the issue I am discussing.)

But to assume so would, in fact, be to confuse the general, or universal, happiness with the greatest possible general, or universal, happiness, and it should be clear—though our discussion of satisficing may help to make it clearer—that these things are not the same. Someone may aim at his *own* happiness

in a satisficing way, i.e., without aiming at his own greatest possible happiness; and if someone's *sole aim in life* were to become a good tennis player, it would hardly follow that she aimed at becoming the best player in the world. (Even the latter aim is a satisficing one, since it need not involve aiming to be as good as possible relative to other players, e.g., in a totally different class from everyone else. Some satisficing aims can thus, somewhat misleadingly, be characterized by the use of typical optimizing concepts like bestness.) By the same token it hardly follows from the fact, if it is one, that morality aims at universal well-being or happiness, that it does so in an optimizing way. Yet a failure to distinguish universal happiness or the general well-being from the greatest possible universal happiness or general well-being is characteristic of the entire utilitarian literature, and we find Sidgwick, for example, constantly running these notions together (as well as identifying the desire for one's own happiness alone, egoism, with the desire for one's own greatest happiness).²⁰

But from what we have seen, it is possible to aim at the general happiness without aiming at the greatest general happiness. And so it is possible to hold a satisficing form of act- (or act-and-motive-)consequentialism consistent with what utilitarians and others have wanted to say about the object of morality. (It is equally possible to maintain a satisficing form of ethical egoism without contradicting or undercutting the most familiar general expressions of that doctrine.) In addition, the most familiar characterizations of (utilitarian) moral motivation also fail to give any preference to optimizing over satisficing forms of consequentialism. For nothing in the idea of impartial, rational, benevolence or of universal sympathy entails a desire for the greatest possible human happiness; and even someone with the *highest degree* of impartial rational benevolence or universal sympathy may not always aim for the greatest happiness or well-being possible, since it is possible for many (even all) individuals to be satisficers about their own well-being and it is hardly clear that the greatest possible benevolence or sympathy towards (sympathetic identification with) them would require us to

²⁰ See Sidgwick, *op. cit.*, e.g., pp. 285–89; also p. 95. The confusion can also be found in Mill's 'Utilitarianism'.

desire their greatest well-being or thus the greatest general well-being of mankind.

It would seem, then, that satisficing (utilitarian) consequentialism cannot be excluded on the basis of those general characterizations of the purpose of morality and of moral motivation that have been used to defend (utilitarian) consequentialism generally and optimizing versions of (such) consequentialism in particular. And our whole previous discussion of satisficing individual rationality and of the satisficing elements in the common-sense morality of benevolence should clear the way to making satisficing (act-)consequentialism seem a genuine alternative to prevalent consequentialist views. However, I have here provided only the crudest sketch of how a plausible version of satisficing act-consequentialism might be formulated and perhaps the ultimate test of the whole notion of (consequentialist) moral satisficing will lie in how appealing more specific and detailed formulations eventually prove to be.

SATISFICING CONSEQUENTIALISM

Michael Slote and Philip Pettit

II—Philip Pettit

I

Consequentialism is usually defined by some formula like: one should always do what has the best consequences.¹ But there is an ambiguity in any edict of this kind and it has haunted consequentialist writing.

The ambiguity concerns the scope of the definite description 'what has the best consequences' *vis-à-vis* the scope of the predicate 'do'—or, *a fortiori*, the deontic operator 'should'. One reading gives the description wider scope, the other narrower.

Take doing something, P, to involve the intentional and direct pursuit of P. Intentional, so far as the goal is anticipated and desired. Direct, so far as the goal gives the agent his main orientation: he does not pursue it by aiming at another target, in the knowledge that thereby he can achieve P.

The wide-scope consequentialist formula prescribes with regard to what has the best consequences, that one should do that. It does not entail that the consequentialist goal should be direct; nor even that it should be intentional. The agent is to take a line that happens, whether or not he is aware of the fact, to yield the desired result.

The narrow-scope formula, on the other hand, supports both entailments. The agent is required to set out intentionally and directly in pursuit of what has the best consequences. He is to let himself be guided *ex ante* by the formula, not just allow his actions to be assessed by it *ex post*.

To pursue the best consequences is, by whatever metric one chooses, to maximize. Since the policy of satisficing is put

¹ See for example Joel J. Kupperman *The Foundations of Morality*, Allen and Unwin, London 1983, page 94. Kupperman is alive to the ambiguity I discuss and has useful things to say on the matter. The ambiguity shows up within the theory of justice in the distinction between utilitarianism considered as providing merely a 'criterion' of the just society and utilitarianism considered as also providing a 'charter'. On this distinction see my *Judging Justice* Routledge and Kegan Paul, London 1980.

forward as an alternative to one of maximizing, it follows that narrow-scope consequentialism must prohibit satisficing. The wide-scope doctrine on the other hand need not do so. There may be good reason why a wide-scope consequentialist—a wide-scope maximizer, if you like—should prescribe a satisficing policy.

Assume that there are cases where maximizing is a technically feasible strategy. I wish to argue, with regard to those cases, for the two following theses.

- A: There are good (if non-conclusive) reasons for any one of which a wide-scope consequentialist might prefer a satisficing to a maximizing policy.
- B: Unless some such reason obtains, satisficing is an irrational policy for anyone, consequentialist or not, to prefer.

The second thesis puts me in conflict with what Michael Slote maintains in 'Satisficing Consequentialism'.² Besides defending the theses, therefore, I shall also try to undermine his position.

II

The notion of satisficing derives from H. A. Simon.³ It directs us to a decision-making strategy of roughly the following form:

- (1) Set an aspiration level such that any option which reaches or surpasses it is good enough.
- (2) Begin to enumerate and evaluate the options on offer.
- (3) Choose the first option which, given the aspiration level, is good enough.

² Slote wishes to argue, not just that unmotivated satisficing is rational, but that the narrow-scope consequentialist ought to drop the reference to 'best' (in favour of 'good enough') and prescribe such a satisficing policy. I ignore this argument since it is undercut by thesis B.

³ See his *Administrative Behaviour*, Third Edition, The Free Press, New York, 1976; 'A Behavioural Model of Rational Choice' *Quarterly Journal of Economics*, vol. 69, 1955; 'Theories of Decision Making in Economics and Behavioural Science' *American Economic Review*, vol. 49, 1959; 'From Substantive to Procedural Rationality' in Spiro J. Latsis, ed., *Methodological Appraisal in Economics*, Cambridge University Press 1976; 'On How to Decide What to Do' *Bell Journal of Economics*, vol. 9, 1978; and 'Rational Decision Making in Business Organizations' The Nobel Foundation, Stockholm 1978. I am exercising some license in my characterisation of the satisficing strategy, since Simon sometimes describes related but distinct policies as satisficing ones too.

The policy described is meant to serve as an alternative to the following, maximizing sort of strategy:

- (1) Enumerate all the options on offer.
- (2) Evaluate each.
- (3) Choose the best.

The main difference between the two kinds of policy is that whereas the maximizer insulates enumeration, evaluation and choice from each other, the satisficer allows them to interact. One evaluates as one enumerates and if a satisfactory option appears, one may choose before either task is complete.

The enumeration involved in each case will spell out the different salient alternatives before the agent.⁴ It may be that to enumerate any alternative one must have looked ahead to enumerate all. In that case satisficing will be distinguished only in respect of evaluation.

As for that evaluation, there are three points to note. The first is that it is not relativized to a particular purpose or aspect. It represents, not what the agent would prefer in this or that context, but what he thinks is better or worse, all things considered.

Secondly, the evaluation of options is supposed, in line with the consequentialist perspective, to be derived from an evaluation of the likely consequences.⁵ This assumption however is compatible with any of a variety of derivations: in the case of uncertain consequences, for example, the derivation may be by the rule of maximin, maximax, maximization of expected utility, or whatever.⁶

Finally, while the evaluation of options yields a preference

⁴ I assume that there is no problem in recognising the features which ought to be ignored in spelling out alternatives, e.g., such features as acting with one's unused limbs in this position or that, acting at approximately 1 p.m. or a micro-second later. Both maximizers and satisficers are taken to have the relevant sense of salience. For the claim that only satisficers can appropriately enumerate alternatives see Alex Michalos 'Rationality between the Maximizers and the Satisficers' in K. F. Schaffner and R. S. Cohen (eds), *PSA 1972*, D. Reidel, Dordrecht-Holland 1974.

⁵ I take the consequences to include the event in which the action itself consists. The case for such an inclusion can be found in Amartya Sen 'Evaluator Relativity and Consequential Evaluation' *Philosophy and Public Affairs*, vol. 12, 1983, pages 128 *et seq.*

⁶ The usual consequentialist derivation is by maximization of expected utility. Simon seems also to endorse this.

ordering which, so far as it goes, must be complete and consistent, it need not yield one which can be represented by a real-valued utility function.⁷ It allows us to speak of maximizing—or satisficing on—preference fulfilment but not necessarily of maximizing utility.

III

Contrary to thesis A, it appears that there can be no good reason for a wide-scope consequentialist to satisfice, or to prescribe satisficing. There is a plausible argument to that effect.

Suppose that someone proposes a reason R for satisficing. If it is to engage the wide-scope consequentialist, then R must be a statement to this effect: that by pursuing a satisficing policy, seeking options with good enough consequences, an agent produces a beneficial side-consequence, C; and that acting so as to produce good enough directly intended consequences, plus the side-consequence C, is acting in a manner likely to bring about the best consequences overall.

But now an opponent of thesis A can argue as follows. If R holds, then trying to do what has good enough consequences will be promoting what has, or is likely to have, the best consequences. If the agent knows this, however, he can recast his policy as one of trying to do what has the best consequences. He can give up pursuing a satisficing policy and, taking account of consequence C as well as the results to which he had previously attended, adopt a strategy of maximization.

In fact, the opponent may urge, this is what any rational agent ought to do. No matter how important the realization of C is, there are bound to be circumstances in which the agent will think that its loss is more than balanced by the achievement of certain other results. In such a setting, satisficing will not produce what has the best aggregate consequences: R will fail. On the other hand, maximizing with a view to all the possible results, including consequence C, will guarantee that upshot. Thus the maximization strategy is, it appears, superior.

Thesis A is not undermined by this argument because one of the premises is false. This is the assumption that if a strategy is

⁷ On the requirements for such representability see H. A. John Green *Consumer Theory*, 2nd edition, Macmillan, London 1976, Chapters 6 & 13.

justified, in crucial part, by a side-consequence, C, then a similar justification is available for the strategy which differs only in casting C as a consequence to be intentionally and directly pursued.

The premise is false because, in Jon Elster's phrase, some justifying consequences are essential by-products.⁸ The standard example is spontaneity. A particular strategy may be justified in part by the fact that the agent who pursues it will achieve spontaneity of character. It does not follow that we can equally justify the variant policy which elevates spontaneity to the status of a consequence intentionally and directly sought. One does not need to be a psychologist to recognize that someone who makes a regular concern of being spontaneous is unlikely to display that characteristic.

IV

Thesis A will be vindicated if we can identify a justifying side-consequence of satisficing which is essentially a by-product of the policy. It must be a consequence that cannot be pursued by an agent in such a way that the justification applies equally to the corresponding strategy of maximization. In fact there are a number of plausible candidates for the role. I will mention two.

One is the consequence that has always been invoked by Simon in support of the policy of satisficing. This is computational ease: the fact that decisions are made without the time and trouble required for enumerating and evaluating all possible options.

It is only reasonable to admit that in some circumstances at least, the gain in computational ease will compensate for any losses involved in satisficing. But is computational ease an essential by-product of satisficing? Can't an agent take direct, intentional account of the difficulty of going through a full enumeration and evaluation of alternatives and, maximizing his overall preference fulfilment, allow himself in appropriate cases to choose one of the options that satisficing might have selected?

⁸ See his *Sour Grapes*, Cambridge University Press, 1983, Chapter 2. Elster does not give my argument for satisficing. He endorses the policy when maximization is not feasible. See pages 14 and 18.

The approach suggested would have the agent raise himself to a level above the enumeration and evaluation of first-order options. Allowing for the benefit of computational ease, he would first enumerate the short-cutting alternatives to that standard strategy and then evaluate the different strategies available. He would maximize at this higher level, opting for the strategy which seemed best. Applying this policy then, he might end up with a first-order option which would also attract a satisficer.

In many areas such an approach will clearly be self-defeating. To attempt to avoid computational costs by computing the benefits of deviating from the ordinary procedure of enumeration and evaluation may push such costs even higher. Computational ease may be attainable then only by forswearing a certain sort of maximizing computation. It may be essentially a by-product of a strategy such as satisficing.⁹

The problem with computational ease derives from a regress. The computational costs of maximizing at any level can only be counted at the next level up. To weigh such costs in a maximizing procedure, therefore, may be to engage further costs that remain unweighed. Satisficing offers an exit from the predicament.

A second side-effect which might be quoted in justification of satisficing is the habit of moderation, the habit of making do with less than the guaranteed optimum, which it encourages. We may ascribe the sort of justification in question to Solon, in view to his association with the maxim 'Nothing too much'.¹⁰

The habit of moderation is certainly something beneficial, as

⁹ This consideration is certainly in the spirit of Simon and is sometimes near to being explicit in his writing. In his Nobel Foundation piece he writes for example 'In long-run equilibrium it might even be the case that choice with dynamically adapting aspiration levels would be equivalent to optimal choice, taking the costs of search into account. But the important thing about the search and satisficing theory is that it showed how choice could actually be made with reasonable amounts of calculation, and using very incomplete information, without the need of performing the impossible—of carrying out this optimizing procedure'. Section 2C.

¹⁰ Richard Routley suggests this kind of justification in 'Maximizing, Satisficing, Satisficing', Discussion Papers in Environmental Philosophy, Research School of Social Sciences, Australian National University, Canberra 1983. I am indebted to his paper for a number of lessons, including the reminder about Solon.

long tradition has it. Lacking a well-tempered character, a person is liable to display obsessive perfectionism, intolerance of the accommodations which life requires, and the like.

But is the habit of moderation essentially a by-product of a strategy such as satisficing? Arguably, yes. Consider the corresponding maximizer who puts a concern with the tempering of his own character into his calculations as to what is after all best. He does not have to raise himself to a level above ordinary enumeration and evaluation like his counterpart in the other case. He simply allows his overall evaluation of the options enumerated to be affected by their significance for moderation of character.

Such a person, for all that has been said, may turn out not to be moderate. On the contrary, he is required to be a perfectionist, if not about moderation, at least about the complex of values, including moderation, which dictates his ranking of options. Given that ranking, he can have nothing but the best; anything less is failure. The direct, intentional pursuit of moderation, under a maximizing brief, is self-defeating. The habit of moderation is an essential by-product. Or so it can plausibly be argued.

There is an important difference between Simon's and Solon's rationale for satisficing. We should comment on it, since the distinction is of relevance later.

If one satisfices for Simon's reason, then one always chooses the best from among the alternatives already enumerated and evaluated. Satisficing on such a ground is intentionally not maximizing but it is not intentionally sub-maximizing, as we might say. One does not seek the best but neither does one seek not-the-best.

Where one satisfices out of Solon's motive, this result is not assured. Suppose that for whatever reason the options enumerated and evaluated include a number which exceed the aspiration level; they may even be an exhaustive set. The evaluation will be based on all relevant considerations, since it is not relativized by aspect or purpose. But these considerations do not include the effect on the agent's habit of moderation. Thus, just as the agent sets himself ordinarily to satisfice in view of that effect, so he may set himself to choose less than the best—so long as it is good enough—in the sort of scenario envisaged.

He may commit himself to a policy of sub-maximization.¹¹

We have seen enough to be able to say that thesis A holds. There are at least two good reasons for either of which a wide-scope consequentialist might prescribe a satisficing policy.

V

We may turn now to thesis B: the claim that satisficing is an irrational policy to prefer, in the absence of a wide-scope consequentialist motive. I will argue that an unmotivated satisficer of the kind envisaged is committed to unmotivated sub-maximization and that this is profoundly irrational.

The unmotivated satisficer prefers a strategy guaranteeing only a good enough result to one ensuring the best. It is not that the enumeration and evaluation required to guarantee the best has costs such as Simon or Solon would count. It is just that the exercise is irrelevant, since a good enough option is preferred to the best.

This means that such a satisficer will have to support a policy of sub-maximization in any situation where all the options have been enumerated and evaluated. He will opt for less than the best, so long as it is good enough. In this he will resemble a certain sort of Solon-satisficer but the difference remains that he can have no reason for the line he takes. He will be an unmotivated sub-maximizer.

Unmotivated sub-maximization is an irrational policy to prescribe or pursue. This fact can be brought out in two ways.

Suppose that an agent ranks A above B, all things considered, and then chooses B without any motive of the Solon type. The irrationality of the policy first appears in the fact that whereas he could have given a reason for choosing A—it is in his view a better option—he can give no reason for choosing B.

It will not do for him to say simply that B is good enough. That might be a reason for him to go for B, if he were unaware of the nature or value of alternatives. It is not a reason for him to choose B rather than A.

¹¹ It is this fact, I believe, which leads Routley, *op. cit.*, to describe the strategy as satisficing rather than satisficing. I define satisficing independently of the rationale offered in support of it. Presumably the follower of Solon will not sub-maximize in the cases under consideration if he is quite sure that his desires, and therefore his overall ranking, are already moderate—and if he is confident that maximizing will not have an adverse effect on his established moderation.

The second way in which the irrationality of the sub-maximization appears is even more striking. To evaluate A as better than B is to be disposed to choose A, other things being equal. For the Solon-satisficer who sub-maximizes other things are not equal. For the unmotivated sub-maximizer however they are. His strategy deprives the notion of evaluation therefore of its usual content. It is not clear what it can mean to rank A above B if when other things are equal one insists on choosing B.

VI

Thesis B is overwhelming, in my view, but it may be useful to pinpoint some reasons why it might be mistakenly rejected.

A crude reason would be that the sub-maximand is assumed to be, not fulfilment of one's preferences, but consumption of the main commodity involved in the options. One sub-maximizes, not in relation to a preference ranking of ice-creams, but rather a ranking of ice-creams by the quantity of the stuff they contain. On this assumption sub-maximization will appear to be so normal that no question may be raised about sub-maximizing without a Solon-like motive.

A slightly less crude reason for rejecting thesis B might be that the sub-maximand is taken to be the fulfilment, not of one's preferences all things considered, but only of one's preferences in a particular dimension. Sub-maximizing on one's aesthetic or hygienic or even moral ranking of options is unexceptionable. Thinking along these lines, a person might slip into thoughtlessly condoning unmotivated sub-maximization.

These reasons are crude because even the sub-maximization that is mistakenly envisaged in each case is not pursued without motive. That has to be overlooked, if the mistakes are to have the effect of making unmotivated sub-maximization seem reasonable. A third reason why such a strategy might be condoned is more sophisticated.

This is the thought that even an unmotivated sub-maximizer has a meta-preference for choosing less than the best, that his choice will fulfil this meta-preference, and that it can therefore be represented as rational. At a level of second-order, preemptive preferences, it guarantees maximal satisfaction.

This sort of justification is circular. The unmotivated sub-maximizer prescribes the fulfilment of the meta-preference

mentioned. But why? Not because that meta-preference exists anyhow, being inherited from nature or whatever. Nor because it is a meta-preference desirable on Solon's grounds. Rather, for the reason that fulfilling that meta-preference just is sub-maximizing. The circle is of small circumference.

VII

Enough has been said in direct and indirect support of our two theses. But the second thesis runs counter to something that Michael Slote maintains and I would now like to make some remarks on his position.

Slote countenances as rational, not just satisficing, but sub-maximization.¹² And not just any sort of sub-maximization, but in particular the kind unmotivated by a wide-scope consequentialist consideration. There is no question of saving computational costs or of practising an ascetic strategy.¹³ One rejects what is best if one can otherwise get what is good enough.

In support of his view Slote argues that unmotivated sub-maximization is countenanced in common sense. You have the opportunity to enjoy a snack but you decide you are happy enough as you are. You have the chance to fulfil any wish for yourself or your family and you ask for a benefit well short of the optimum. You are in a position to offer food and hospitality to some people in need and you give less than the best. In each case he maintains that common sense will respect your decision as rational.¹⁴

The last example is best ignored. What is morally best, even if it seems to demand no special sacrifice, may not be that which the person regards as best all things considered. The final ranking may be affected by all sorts of non-moral considerations, such as the wish not to have to think too well of oneself or not to have to appear too saintly to others. What common sense applauds in this case is not clearly an instance of sub-maximization.

Of the other examples, the first is easily handled. You are

¹² See page 146.

¹³ See pages 145 and 146.

¹⁴ See pages 143–149. With the first two examples Slote appeals to what we would intuitively say, with the third to what the common sense morality of benevolence would prescribe. I have cast both as appeals to what common sense would say.

supposed not to have any ascetic motive, such as Solon would provide, for choosing less than what you regard as best. You are faced with a choice between having or forgoing a snack. You forgo it. What ground can there be then for saying that you nevertheless regarded having the snack as, all things considered, the better option? You may have realized, as Slote says, that you would enjoy the snack but clearly you still came to the conclusion that it was better to practise restraint. There is no evidence here of unmotivated sub-maximization.

Slote more or less concedes the case. 'Although he knows he would enjoy the snack, the very fact that he rejects such enjoyment might easily be taken as evidence that he doesn't in the circumstances regard such enjoyment as a good thing'¹⁵. Conceding this, he directs our attention to the remaining example.

The case is a fairy-tale one. 'The hero or heroine, offered a single wish, asks for a pot of gold, for a million (1900) dollars, or, simply, for (enough money to enable) his family and himself to be comfortably well off for the rest of their lives. . . . Someone who makes such a wish clearly acknowledges the possibility of being better off and yet chooses—knowingly and in some sense deliberately chooses—a lesser but personally satisfying degree of well-being'.¹⁶

In this case, as in the others, there is every reason to protest that the behaviour is not self-evidently an instance of unmotivated sub-maximization. The agent may be superstitious about asking for more. Or he may think it is wrong to use such an opportunity to become more than moderately rich. He may believe, all things considered, that it is best to be financially at the mean, or only just above it. Or he may even be infected by the wisdom of Solon. The possibilities are legion.

It should be clear that Slote's recourse to common sense is not likely to yield the conclusion he desires. We may concede that what common sense deems to be rational can hardly be denied that title. We may agree that common sense will judge the cases in question as he says it will. But we can always show, for every case presented, that it need not be taken, and may not be taken by common sense, to illustrate unmotivated sub-maximization. Against such a line of attack, thesis B will easily stand firm.

¹⁵ Pages 146-7.

¹⁶ Page 147.

VIII

In conclusion, an *ad hominem* point against Slote. This is that the sub-maximizing strategy which he countenances imposes the same costs as maximizing, but without offering the same benefits. It gives us the worst of both worlds.

Slote implies that in order to sub-maximize, an agent needs to have enumerated and evaluated all the relevant options. What he endorses as rational therefore is a policy of mimicking the strategy of maximization up to the moment of choice. One enumerates all the salient alternatives; one evaluates each; and then one chooses something less than the best, provided it is good enough.

Slote commits himself to full-scale enumeration and evaluation of alternatives, because he maintains a maximizer's conception of how to evaluate options. He takes the measure of what is sufficiently good to be the distance from what is the best available; he takes sufficient goodness to be judged by a comparative metric.¹⁷ If one has to compare an option with all the alternatives, in particular with the best, before one can tell that it is sufficient, then one must enumerate and evaluate all those alternatives before making one's choice.

The strategy which he countenances is a rare and perverse mixture. It requires that the agent should suffer all the pains of maximizing just in order to ensure that he does not enjoy its fruits. Such a strategy hardly constitutes a serious challenge to our thesis B.

NOTE

For useful comments and conversations I am indebted to a number of colleagues, some resident and some visiting, at the Research School of Social Sciences, Australian National University. I am particularly indebted to R. E. Goodin.

¹⁷ See pages 154 and 156-7. Slote's emphasis on the comparative accounting of options is an endorsement of the economist's notion of opportunity costs.