26. The reasoning in this paragraph follows some reasoning of Shelly Kagan, *The Limits of Morality*, Oxford University Press, 1989, 371–4, where Kagan is investigating how a defender of agent-centered options can avoid moral decisiveness. His point is that it is hard to see how an option can avoid turning out to be a requirement.

27. Alastair Norcross and Frances Howard-Snyder, "A Consequentialist Case for Rejecting the Right," *Journal of Philosophical Research* 18 (1993): 109–25.

28. Here I have to add that I am not claiming to have solved the problem that Kagan raised in the pages of *The Limits of Morality* cited above. I am claiming to have solved a related problem, inspired by Kagan's discussion. Kagan's actual problem is harder to solve, because it amounts to providing a thorough justification for options rather than removing the air of incoherence that surrounds them when viewed in a certain light. I think his demand would be satisfied by an explanation, mentioned but not offered in my text, for why and in what sense considerations of duty should be stronger or weightier all things considered than considerations of beneficence. But as I say in the text, I can only speculate about the existence of an explanation like that. I don't have one.

8

The Plausibility of Satisficing and the Role of Good in Ordinary Thought

Mark van Roojen

When we think about whether it ever makes sense to choose something that is simply good enough even when other better things might instead be chosen, it seems that it can. We can readily be offered examples in which it seems that is just the choice we sensibly make. To use one oft-cited example,[1] it makes sense to accept a reasonably good offer on a house one is selling rather than hold out for a higher price. It may be just as reasonable to hold out for more, but provided the offer is good enough, there is nothing irrational or unreasonable in accepting the first sufficiently good offer.

Thus the examples seem to show that satisficing – that is, to choose the merely good enough over an option which is better yet – is sometimes rational. However, many philosophers have wanted to argue that things are not as they seem. They wish to defend the idea that satisficing is rational only if it serves as part of an overall strategy to maximize. In service of this position, they have available a general strategy for dealing with examples that purport to show otherwise. This strategy is to ask the advocate of satisficing what it is about the lesser option that justifies one's choosing it over the greater. Once a reason is offered, the clever proponent of maximizing can incorporate that consideration into a more sophisticated characterization of goodness, so that options which satisfy the consideration will, other things equal, be better than alternatives. Then, using this more sophisticated way of constructing a notion of goodness, the defender of maximizing can argue that the supposedly merely satisfactory option is in fact better than the alternative.

Thus, in the foregoing example of selling the house, an advocate of maximizing can argue that by taking the first sufficiently generous offer,

the seller really is maximizing the good, although perhaps not maximizing the selling price. The reasons why it makes sense to take the first good offer include the fact that one cannot be certain of getting a better offer, that one will have to spend more time waiting for a better offer if it is to come, that the uncertainty during that time will create some anxiety, and that such anxiety is not worth the extra money that the seller might get. Thus, the advocate of maximizing the good might argue, the seller is maximizing expected good if not money.

In this paper, I want to resist this sort of argument while granting it its due. I believe that the defenders of maximizing are right to ask for reasons to choose an option that is admittedly less than the best. And, the formal trick of constructing an ordering incorporating the very features that ground the reason offered in defense of maximizing can also be accomplished. So the defense will have to proceed by arguing that not every reason can be given its *best* characterization when presented in the form of placing an option in an ordering from better to worse. My underlying idea is that we really need two distinct value notions, one roughly captured by the word 'right' as it is ordinarily used, and the other involving a notion of goodness. Because the strategy that assimilates satisficing choices to maximizing a more sophisticated conception of goodness tends to push these two notions together, it has costs that are not worth paying for the sake of theoretical simplicity.

Thus I will argue for the following nest of theses: Satisficing in conditions when one is not maximizing can be defended as rational, provided that nonconsequentialism is rational and provided that the preferred characterization of the resulting nonconsequentialist position is not one in which the right action is justified in virtue of maximizing agent-relative value. Rather, the nonconsequentialism that can serve to defend satisficing should be one in which the best characterization of certain reasons to act does not involve maximization of value of any sort, whether agent-relative or agent-neutral. I will also argue that there are reasons to prefer this sort of nonconsequentialism to theories that defend the sorts of actions distinctive of nonconsequentialism by claiming that value should be thought of as agent-relative value. An upshot of the argument will be that satisficing cannot be well defended within an overall framework that is consequentialist.

Some Clarification of the Subject Matter

My topic is not about the *self-regarding* rationality of satisficing or maximizing. I take it that by delimiting the issue by using the term 'self-regarding'

one has restricted the reasons to maximize or satisfice to reasons that have to do with promoting the interests of the agent in question. But once we've done that, I don't see how we can resist the conclusion that satisficing is rational only as a strategy to maximize expected personal good. And if we limit ourselves to self-regarding reasons, then any strategy that conflicts with maximization will be criticizable for not securing as much personal good as possible; and the proponent of mere satisficing will not be able to point to a competing *self-regarding* consideration which outweighs that demerit, for all of the self-regarding considerations will have already been weighed up in ranking the various options from the perspective of self-regarding rationality.

My topic is rationality without such qualifications. I will assume that the overall rationality of a choice is a function of all the considerations that can be brought to bear against and in favor of each of the options. These will often include considerations such as how well the option serves the agent's own interests, but they may also include altruistic considerations, moral considerations that are neither purely altruistic nor self-interested, and, for that matter, aesthetic considerations when they are relevant. This will allow the issue that I consider to be most basic to emerge in its most natural way: Are there considerations for or against choosing certain options that are not best captured by representing those options as maximizing some good or other?

Thus I want to defend the idea that it can indeed be rational to knowingly choose an option that is less good than the alternatives. To label the position, I will be arguing that mere satisficing can be rational, provided that nonconsequentialism can be vindicated. I do not claim that it is always rational nor even that it is rational most of the time. To clarify, what I mean by "mere satisficing" is choosing an option that is merely good enough but not the best, where that choice should not be reconstrued as part of a sophisticated strategy for maximizing overall goodness by satisficing local goodness,[2] nor for maximizing expected goodness by taking an outcome one knows to be good enough when the costs of continuing to search for a better option are high.

But I don't want to defend the claim that it can be rational to do a worse action when one knows that by doing something else one would be doing a better action.[3] When "better" is applied to actions, it no longer seems to me to be about the goodness of the option chosen, but rather about the rightness or rationality of the action. That we should rationally do what is most rational is a tautology that proponents of mere satisficing should not be portrayed as denying. The rationality of mere satisficing is thus not best construed as requiring that one "may be rationally permitted to

do less than the rationally ideal or best."[4] If it is most rational to choose some particular option, then it is less rational to chose any other option.[5] Once we have ordered a series of choices by their rationality, there is no substantive issue remaining. We may decide to call any of the set of choices toward the more rational end 'rational' or we may reserve that name for only the most rational choice. But nothing of any substance can turn on this. There may be a substantive issue about whether we should criticize choices that are close enough to the most rational choice. And we could align our use of the term 'rational' with our verdicts about this issue. But in this case the substantive issue is about *us*, and not about the rationality of the agent's choices. The substantive question concerning rational satisficing is whether a choice can be as rational as it gets without optimizing some good or other defined independently of the rationality of the action itself.[6]

I should also probably stress that the issue I am pursuing is about satisficing *goodness* and not about satisficing whatever is represented according to a theory of revealed preference. It would be incoherent for a theory to define utility as whatever was maximized by a person's actual choices or dispositions to choose and then go on to say that one should act so as to do something other than maximize whatever this was. How could one follow such advice? If one did so, one would be making different choices and manifesting different dispositions to choose, and this would change one's preference ordering as conceived by the theory of revealed preference. The theory would then have to use these preferences as the basis for its advice and on that basis advise us to do something still different, and if we follow the revised advice, the theory would run into the same problem all over again.[7] For the position to make sense, the thing that we are satisficing has to be something that is not solely a function of our preferences as revealed by our choices. Because I think that there is something like objective goodness independent of our preferences that allows us to compare options, this is the notion I want to use in asking and answering the question at issue.

Enough about what I'm not defending.

A Strategy for Defending the Rational Mandatoriness of Maximizing

The kinds of situations that interest me here are those in which a person chooses an option that is good enough but that is less good than available alternatives while still believing that those better alternatives are available. Defenders of mere satisficing have proposed various examples,

and, as I have already remarked, those who think that mere satisficing is irrational have a general strategy for combating arguments employing those examples. They can ask the defenders of mere satisficing to offer reasons in favor of choosing the lesser but adequate option and then use those reasons to argue that the option chosen is not after all the lesser of those available. If the advocate of mere satisficing refuses to explain why it makes sense to choose the lesser option, then the critics can offer considerations of their own and employ them to show that the option chosen is really best, or they can refuse to agree that the choice makes sense. If their reaction is the latter, who could blame them? If theorists want to argue that a choice makes sense, then they should have something to say to doubters. At least if one is a doubter and the advocates of the view you doubt don't say anything to defend it, one hasn't been given any reason to change one's mind.[8]

In the example we began with, that of selling a house, it seems relatively easy to come up with reasons for taking the first satisfactory offer. If we don't take the first satisfactory offer, one thing we will lose is the time it takes to get a better one. In general, money now is worth somewhat more money later; that's why people can make money in the loan business. By taking the money offered right now, we can immediately put it to good use, perhaps in making more money, or buying something we need or want. We also forgo the trouble and stress of dealing with additional potential buyers, realtors, and so forth in the attempt to gain still more. These reasons, which certainly count in favor of a satisficing monetary strategy, in turn give us reasons to think that we might actually be maximizing the overall good, composed of money and enjoyment and other non-monetary goods. Furthermore, there is always some chance that we will not later get a higher offer or even another offer as good, in which case this offer is the best we are going to get even in monetary terms alone. Even where that is not actually true, antecedent to knowing how things will turn out, accepting the offer may maximize expected monetary benefit.[9]

A somewhat more challenging example is suggested by Michael Slote. It may, he argues, make sense for someone to aim "to be a really fine lawyer like her mother," rather than to choose to be the best lawyer she could be.[10] So far, this does seem to capture the judgment of ordinary common sense. But this may be so only because excellent lawyering can interfere with other worthwhile pursuits, and because taken to an extreme it can interfere with enough of them to make one's life less good as a whole. If so, the example presents no obstacle to the view that suboptimal choices

are less than fully rational. The woman in question can be depicted as choosing the best life for her, which is not the life in which she is the best lawyer she could be.

But Slote goes on to deny that this is what is at work in the example. He claims that common sense will endorse the woman's choice even when it is not made "from a belief that too much devotion to the law would damage other, more important parts of one's life."[11] And he mentions moderation as a trait that might explain such choices and help us to regard them as both nonmaximizing and rational. Once again, there may be reasons why moderation itself can contribute to the greater good, so that the choices can be reconstrued as optimizing after all. Moderate goals may just be more realistic and give agents the greatest chance of success. Or, small steps may be the best way to do as well as one can. Thus, aiming to be a really fine lawyer may be the best way to become the best lawyer that one can be.[12]

But the fans of satisficing can and do respond by citing the intrinsic value of moderation itself or of other virtues that satisficing or moderation serves. Perhaps certain suboptimal choices are the manifestation of dispositions that are involved in certain virtues. If it can be rational to value those virtues, then satisficing will be rational. This is the strategy employed by Christine Swanton. She argues:

> [T]he rationality of satisficing stems from the value of acting from desirable or at least not undesirable traits. It is characteristically rational to act out of friendship, love, courage, and so on, even where such action does not directly or indirectly optimize. This rationality stems from the rational desire to be the sort of person who is a friend and acts as one, who expresses his love for someone, and so forth. Such a person will not always set aside those traits and emotions in order to optimize . . . where the cost is betraying, or not being true to, his character, his love, or even his feelings. Let us assume that the agent realizes that the optimific action is to betray his friend. But the agent does not want to be the sort of person who betrays his friend to produce a better state of affairs. . . .[13]

This is the hardest challenge for the foes of mere satisficing to meet, and in the end I will come around to defend something akin to Swanton's position. But the foes of mere satisficing will not be easily persuaded by the value of virtue, precisely because they think that this value should be added in as one of the factors determining what is better than what. With the added value of virtue in the mix, choices that seemed not to maximize value can now be reconstrued as doing so after all.

Having admitted virtue in as one of the intrinsically valuable things that makes an outcome better, there are two ways that the friend in the example may be optimizing even when he seems to be refusing to do so. First of all, as Aristotle tells us, virtues are created and destroyed by acting as though one has the virtue. Thus, if friendship is valuable, and betraying one's friend is not an action typical of friendship, then there is some likelihood that throwing the friend over would undermine the agent's character so that he would be less likely to form true friendships in the future. Thus, one future effect of betraying a friend is that one will have fewer friends over the long haul. This result would be worse than sticking with a friend when it is disadvantageous.[14]

More importantly, the value of friendship itself – and its contribution to one's own state of character – might be so great as to wipe out the disadvantage of sticking by one's friend. Swanton begins her exposition by saying that it is the value of acting from desirable traits that rationalizes the suboptimal choice. But if by this she means that acting from virtue is so valuable that it often justifies giving up other valuable things to do it, foes of mere satisficing will agree. However, they will insist that this is so because that value has to be weighed against the disvalue of continuing to manifest the trait. Apart from the intrinsic value of acting on the trait, the virtuous action does not bring about the best available outcome, whereas when we factor in this intrinsic value the balance tips back the other way. Virtuous action does, in the relevant cases, rationalize acting in a way that is otherwise less than optimal. But once the value of virtue is figured in, such choices are nonetheless optimizing.[15]

I hope that by this point the strategy of argument against mere satisficing is reasonably clear: For any consideration that can be cited as rationalizing a choice that is not optimal, use that consideration to argue that the end chosen is in fact better than the other alternatives after all. If the advocates of rational satisficing are to prevail, they will need a way of blocking the strategy. They will need an argument to show that not all considerations are best captured by factoring them into a story about what makes outcomes better. In effect, what is needed is a way of making the considerations that allow us to order choices in terms of their rationality part company from the factors that make outcomes better. To put it yet another way, we need a theoretical reason to resist reducing the right to the good or vice versa. Thus, it may be helpful to look at a longstanding debate that is often put in just those terms – that is, the normative ethical controversy about the truth of consequentialism. I'll begin by rehearsing an argument from that debate that in some ways parallels the arguments concerning satisficing. I will then offer a strategy by which nonconsequentialists can resist the arguments that parallel the foregoing

anti-satisficing arguments. Finally, I will go on to show how those same arguments can be used to defend mere satisficing.

A Parallel Argument for Defending Consequentialism Against Nonconsequentialism

There is a long history of debate between theories that are intuitively consequentialist – that is, theories which evaluate the rightness of actions by looking at their consequences broadly conceived – and theories that are nonconsequentialist and therefore deny that actions can always be judged right or wrong solely in virtue of their consequences. Since Elizabeth Anscombe first introduced the term,[16] that debate has carried on in part by allowing the consequentialists to characterize consequences of actions very broadly, so that features of actions that were not relevant according to more traditional forms of consequentialism such as utilitarianism are counted as consequences of the action – for example, that a right is violated or that a lie is told will count as a consequence of violating someone's rights or of telling a lie. Consequentialists then argue that given enough latitude in characterizing the consequences of actions, they can show that right actions always bring about better consequences than their alternatives,[17] or on satisficing models of consequentialism that their consequences are good enough. Nonconsequentialists argue that even with the additional resources, consequentialism is not adequate to capture the correct view about which actions are right in key instances because there exist other considerations besides the goodness of outcomes that determine what sorts of actions are right. Generally, nonconsequentialists defend three sorts of considerations that justify actions which are not reducible to the value of their consequences. One sort involves special obligations toward others that trump considerations of overall good. Another involves constraints on permissible ways of bringing about good outcomes. And a third involves the moral permissibility of pursuing one's own commitments even when they would have to be abandoned to pursue the overall good.[18] Nonconsequentialists often argue by presenting examples of actions that they hope their consequentialist opponents will agree are right and that they think cannot be accommodated within the consequentialist framework.[19]

Consequentialists respond by bringing to bear various consequences that their opponents may have left out of account, so that the action which intuitively seems right turns out to have better consequences than the alternatives after all. For example, in explaining the rightness of refusing to violate the rights of one person to save a greater number of others, they argue that harms caused by rights violations themselves have a greater negative weight than the same sorts of harms when they are not the result of rights violations. These claims have some plausibility, and they work to defend the intuitive response to the cases at hand, so long as the harms prevented by violating the right in question do not themselves involve the violation of similar rights.

Eventually however, the debate reaches examples in which capturing the intuition that nonconsequentialists expect us to share requires bringing in agent-relative conceptions of better or worse in order to remain consistent with the idea that right actions have better outcomes than their alternatives that are not right. One sort of relevant example involves special obligations that we may have toward those close to us. Intuitively we think it reasonable for parents to choose to save their own children from among a group of children who are in danger of drowning, even if it might be the case that they are better placed to save some other child or several other children. At least at first, it is not obvious how consequentialism can accommodate this point. Isn't it obvious that saving two children brings about a better outcome than saving just one?[20] At this point, advocates of consequentialism such as David Sosa present arguments like the following:

> To the extent that we believe the father should save his own son, we reflect a preference for the set of consequences available to the father that includes his saving his own son. Although very similar to each other, the sets of consequences are not equally son-savings. The total state of affairs consequent on an act, and the value of that total state of affairs, will depend on who performs the act. In this way consequentialism can be an agent-relative ethical theory.[21]

Sosa is actually wrong that what he has said so far involves agent-relative values. Agent-relative values allow different agents to view the very same state of affairs as having different values. Although Sosa's suggestion values a relational property of actions (son-saving), it does not give this property agent-relative value, for all alike are required to view the action as valuable for instantiating this property.

However, to fully capture the intuitive thought about a parent's obligations, Sosa is right that we will need genuine agent-relative values. The father in Sosa's example saves his own child not because he cares about who does the saving but because he wants *his child* to be saved. He would just as much want someone else to save his child as he wants himself to save the child, as is shown by the fact that he wouldn't elbow the

lifeguard out of the way should a lifeguard be about to save the child.[22] Thus we need to allow that the best consequence for parents in the situation to be one where *their own* child is saved, if we want both to endorse the intuitive thought about the example and at the same time to save the consequentialist thesis that the rightness of actions is always a function of the goodness of the actions' outcomes. Thus, some advocates of consequentialism urge a liberal interpretation of what constitutes good consequences, one that allows the relative goodness of consequences to vary with the evaluator's perspective.

Another sort of idea favored by nonconsequentialists will also require agent-relative values, and perhaps also time-relativity. This is the notion that there are side constraints on right action – that there are certain action types which should not be done, even to achieve what all agree are benefits. Candidates include prohibitions on the killing of innocents, torture, certain sorts of dishonesty, and so on. These constraints also require agent-relative betterness orderings to remain consistent with the idea that right actions always bring about the best outcomes, for advocates of side constraints insist that it would be wrong to violate such constraints even to prevent someone else from violating the very same constraint. And, on one reading of what side constraints require, mere *agent-relativity* won't fully capture their stringency. If such constraints require a person not to act dishonestly, even when that person knows that honesty now will lead to his or her own greater dishonesty later, then one will need time-relative orderings to do the trick, ones that allow current actions to weigh more heavily than future actions.[23]

Thus some consequentialists believe that they can defend their theory in the face of counterexamples by admitting such relativized values. And many who are sympathetic to consequentialism are willing to stipulate that the addition of such values no longer makes the view consequentialist strictly speaking but go on to defend the resulting theories that invoke agent-relative or time-relative values as more adequate than more standard nonconsequentialist alternatives.

Reasons to Resist the Argument
Where Consequentialism Is Concerned

A problem with this strategy of argument as a defense of consequentialism is that it is too powerful. Once we allow ourselves to individuate consequences or states of affairs in such a way that they are distinct whenever a reason can be offered to prefer one to the other, then the theory

that we rationally must maximize good or expected good has no content anymore. Any choice can be described as bringing about the best consequences or expected consequences. This is the objective analogue of a similar problem for theories that require us to maximize expected subjective utility. On the one hand, if we do not allow ourselves such fineness of grain, seemingly reasonable choices will be ruled out by the theory. On the other hand, if we are allowed to individuate options in a sufficiently fine-grained way (apples when oranges are the alternative rather than just apples), then *any* set of choices can be made consistent with the theory. The theory no longer gives us any normative advice.

The problem is not quite as bad when we are in the realm of ranking options by their objective goodness as opposed to dealing with preferences. We have required that the options be individuated in such a way that they count as different options if some reason can be given for choosing between them. Thus some options that are intuitively the same will be treated as the same option by the theory because there will be nothing to choose between them. So the theory is not quite consistent with just any set of choices. But because more theories will now be characterized as consequentialist or consequence-based, less is ruled out by this sort of consequentialism. And partly because of this we still have to pay some costs.

One cost of this way of defending consequentialism is that it causes us to lose track of the ongoing debate between consequentialism and nonconsequentialism. If everyone can be characterized as engaging in consequence-based evaluation, what were the arguments about? We can recover the difference between the positions if we remember that old-fashioned consequentialists did not countenance agent-relative values, whereas a number of distinctively nonconsequentialist notions such as side constraints can be captured in a consequence-based framework only by using agent-relative or time-relative features of outcomes to individuate and rank them. Thus we can say that the debate we were engaged in was really a debate between the advocates of agent-neutral and time-neutral rankings of states of affairs and those who wanted to introduce agent- and time-relativity into the rankings. We can at least describe the positions in a way that shows that they are distinct and carry on our discussion of which sort is more likely to be correct.

But I suspect that our discussion will have lost something, at least from the point of view of a nonconsequentialist. Many reasons for distinctively nonconsequentialist choices do not present themselves as grounded in agent-relative or time-relative values. What makes the

nonconsequentialist choice to honor a side constraint make sense to the agent is not that it maximizes agent-relative value to do so, though once we see that it makes sense we can construct an agent-relative ranking of the options in terms of their choiceworthiness. The reason it makes sense to honor a constraint may be, for instance, that this right which would be violated is important, and the agent-neutral benefits of violating the right are not such as to swamp that consideration.

The point for my purposes is this: Although agent-neutral evaluations of goodness and badness play a role in even most nonconsequentialist deliberation about what to do, agent-relative rankings of options don't play the same role in nonconsequentialist deliberation. Rather, the considerations that allow us to construct the agent-relative orderings present themselves in a more direct fashion, as considerations that compete with the agent-neutral goodness of an option for our attention. To use the familiar taxonomy of reasons, prerogatives that allow us to pursue special projects not conducive to agent-neutral goodness, deontological constraints on certain ways of promoting goodness, and special obligations due to special relationships all interact with the goodness of outcomes to determine what we should do. When we capture our normative theory in one agent-relative ordering that is itself a function of both agent-neutral value and these sorts of reasons, we obscure the ways in which these various kinds of reasons interact. We have lost information. I'll illustrate before going on to argue that we want a theory of rightness to generate the sort of illumination that is lost when we agglomerate all the relevant considerations into one agent-relative ordering of states of affairs from better to worse.

Recall the example of a parent whose child is among a number of children at risk. I will suppose that common sense dictates that parents are right to aim to save their own children, even when they can do more to save other children. But this does not require a judgment on our part that the state of affairs which is a son-saving is better than the one in which a parent saves another child. We might have no reason to prefer the one over the other, or even a reason for preferring that the unrelated drowning person be saved. Perhaps the talents of the unrelated swimmer are more likely to benefit humanity. Thus, if the father in Sosa's example wrongly decided to save an unrelated child, we would do nothing to redirect his energies.[24]

Nor does the rightness of the action require the agent – the father in Sosa's example – to make such a judgment. Even while viewing the saving of his son as the right thing to attempt, a father can recognize the legitimacy of our outsider's view. Just as we can allow that he did the right thing in saving his son though we wish that he had done otherwise, he can allow that from our point of view things might have been better had they gone otherwise. He might even agree that it would have been better overall if he had saved the other swimmer, because of that swimmer's great ability to benefit humanity. And recognition of such agent-neutral goodness and badness is not always idle in determining the right course of action even within a nonconsequentialist point of view. This is the important point to see in grasping the inadequacy of theories that reduce rightness to choosing the action that is best in an agent-relative betterness ordering. A great objective need on the part of some other child might delay the father's mission to help his own child. Enough agent-neutral goodness obtained even at some risk to his particular obligations to those close to him might legitimate his modifying his plans. Thus, the agent-neutral goodness of an outcome can sometimes override distinctively nonconsequentialist reasons to determine the rightness of an overall course of action, at least on the nonconsequentialist theories that get the most support from common sense.

The agent-neutral goodness or badness of a resulting state of affairs can also defeat the force of other reasons that might normally justify not bringing about a neutrally better outcome. Side constraints against doing certain act types can be overridden if the only way to prevent serious harms is to do those actions, or perhaps even if doing those actions were to bring about enough benefits. The same is likely true of agent-centered prerogatives to pursue one's own projects and goals. Whatever one is up to, certain sorts of emergency can suggest that we should put our private goals aside to prevent disaster.

Agent-neutral goodness interacts with at least some of these reasons in another way. In deciding between various courses of action required by traditional nonconsequentialist reasons, the agent-neutral goodness or badness of the various options may play a role. For example, although we might let our special obligations allow us to ignore goods and bads outside of our realm of responsibility so long as they are not overwhelming, within our realm of responsibility the goodness and badness of various outcomes may determine our choices. When various family members have competing needs, we can use the importance of those needs – a perfectly agent-neutral importance – to decide how to act. Thus from this sort of nonconsequentialist perspective, agent-neutral goodness along with a variety of other reasons to act is playing more than one role in determining what we should do.

I claim that it is much more illuminating for a moral theory to highlight these roles played by agent-neutral goodness than to fold it in along with other considerations to generate one agent-relative ranking of states of affairs from worse to better. A theory that highlights the distinct grounds of our reasons for acting in different circumstances does more to rationalize what we are up to. If asked why we are doing what we are doing, we might have to talk both about reasons grounded in the goodness of the states of affairs that would result if we did this or if we did that *and* the reasons that make it the case that only an agent-relative ranking of states of affairs will allow a consequence-based theory to capture rightness (if nonconsequentialism is correct). What exactly we would need to talk about would depend on the choice at hand and on the various alternative actions. But if someone were truly interested in why we did what we did, or why we thought doing some action or other was right, then this sort of explanation would be more illuminating than one which pointed out that the chosen alternative ranked highest in our agent-relative ordering of states of affairs from better to worse. More to the point, such an explanation would also be more illuminating than an actual enumeration of that ordering which, given the infinite variety of possible states of affairs, we would never be able to give. If we could grasp that ordering, we might by testing various hypotheses be able to work out what features of the states of affairs in question determined the ordering. In that case we could generate the reasons that determine that ordering. Then we would have some enlightenment about our reasons for acting as we do or as we should. But in fact it is the story about the features that generate the ordering that tells us what our reasons are, not the ordering in which those features are not factored out.

Not every true description of options involved in a given choice will serve as part of a rationalizing explanation of that choice. Not just any way of picking out the rationally preferred option shows that it makes sense to choose it. For example, it might help rationalize my choice of a given candy that it tastes sweet. It may also be that what it is for a thing to taste sweet is for it to stimulate certain receptors on a normal person's tongue in a certain way. But an explanation that used this feature of the candy to explain why I chose it would be less than illuminating, at least unless I knew that these were the receptors that constituted the basis for our ability to perceive sweetness. Insofar as showing that an action is right is part of showing that it is rational or that it makes sense to choose it, the story we tell about what makes actions right needs to be sensitive to this. A good account of what is right should be illuminating. It should show

what it is about the action that makes it choiceworthy for an agent. This means, I believe, that a good account should display the option in such a way that it could make sense for a rational agent to choose it when it is thought of in the way employed by the theory.

Frank Jackson, no friend of nonconsequentialism, puts the point nicely in arguing that consequentialism should adopt the *expectation* of goodness rather than the actual goodness as the standard of right action:

[W]e are dealing with an *ethical* theory when we deal with consequentialism, a theory about *action*, about what to *do*. In consequence we have to see consequentialism as containing as a constitutive part prescriptions for action. Now, the fact that an action has in fact the best consequences may be a matter which is obscure to an agent.... Hence, the fact that a course of action would have the best results is not in itself a guide to action, for a guide to action must be in some appropriate sense present to the agent's mind. We need, if you like, a story from the inside of an agent to be part of any theory which is properly a theory in ethics....[25]

Jackson's last point is the crucial one. Insofar as an agent-relative ordering of all states of affairs from better to worse can be constructed only once we have figured out which actions are right, it cannot serve as a guide to rightness. If ethical theory is to depict what is going on in such a way that it represents decision making "from the inside," it will have to give us something more. If nonconsequentialism of the traditional sort presents us with the right story about what makes actions right, then a theory that keeps agent-neutral goodness apart from the sorts of considerations that justify failing always to aim at such goodness will be superior. Thus, if the various strategies to argue for nonconsequentialism succeed, a theory that puts some distance between rightness and goodness is to be preferred to one that assimilates the latter to the former.[26]

Deploying the Nonconsequentialist Resistance Argument to Defend Mere Satisficing

It should be evident that if the foregoing is correct and if nonconsequentialism is the right account of morality, optimizing goodness is not always morally required. But of course, this by itself does not show that satisficing goodness makes sense. First off, we would need a defense of nonconsequentialism, something I have not offered. Bracketing that caveat, even if nonconsequentialism were correct, not every form of choice that eschews optimizing is a form of satisficing. To satisfice, one must pay attention to the level of goodness without aiming to maximize goodness above some satisfactory level. Thus a moral theory that postulates absolute side

constraints will not in virtue of those constraints rationalize satisficing, even if it requires pursuing the good when those side constraints are not violated, for on such a theory the thought that an outcome is good enough has no real role to play in rationalizing actions. Rationalization will come either from the applicability of the side constraints in question or from the fact that the option is as good as possible without violating those constraints.[27]

A nonconsequentialist theory with defeasible side constraints will, however, make sense of satisficing. The sort of theory I have in mind allows that if things get bad enough, then the normal side constraints on right action can be overridden by duties to help others. Different theories of this sort have been proposed by Thomas Nagel and by Michael Walzer in discussions of justice in wartime.[28] The idea is that normally there are constraints on what we may do to other people that prevent us from aiming to kill innocent people, torture, and so on. But if the alternative to doing so is bad enough – which is to say not satisfactory – then we may violate these constraints to avoid serious calamities. This sort of theory allows the acceptability of refusing to violate the relevant constraint to play a role in rationalizing the choice, for if the state were not at least acceptable, we would not have made that choice. This sort of theory makes sense of satisficing goodness conceived of in the agent-neutral way that I have argued even nonconsequentialists have a reason to recognize. A similar sort of defeasibility might obtain for permission to pursue one's own projects and the special obligations we might have to others. Thus, on the versions of nonconsequentialism that allow anti-consequentialist considerations themselves to be defeated in the face of importantly good or bad consequences, it will make sense for an agent to satisfice. For on such theories, as long as the action chosen has an outcome that is good enough, a person is justified by these other considerations in acting in a way that brings about an outcome that is not the best.[29]

Conclusion

If this is the best way to resist the optimizing strategy of folding all considerations into our characterization of the good of an outcome in hopes of showing that even when we seem to be satisficing we are actually maximizing, then the defense of mere satisficing requires a defense of nonconsequentialism, in fact of nonconsequentialism of a specific sort. Although any reader will likely see that this is the sort of view I favor, I have done nothing to argue for that nonconsequentialism here. So I have not shown that mere satisficing makes sense. I have shown that we can hope to make sense of it only if this sort of nonconsequentialism is correct.

Still, this in itself might indirectly provide some support for nonconsequentialism, at least for those who feel that mere satisficing can make sense. It has been suggested that satisficing consequentialism will be better able to capture the considered judgments of common sense, thereby lessening the challenge that such judgments pose to consequentialism.[30] If my defense of mere satisficing is the best available, mere satisficing will be rational only if nonconsequentialism gives a correct account of our moral reasons to act. If so, satisficing consequentialism is not a viable theory. It will suffer from the same sort of instability that rule consequentialism is often said to suffer from. Although it may tend to do better at agreeing with the judgments of common sense about what we should do, it does so at the cost of undermining its own rationale, for to be a consequentialist is to accept that the rightness of an action is a function only of the goodness of the outcomes produced. Once one has said that much, one has accepted that all reasons are such only by reference to the goodness of a state of affairs in which they are obeyed. Thus, a consequentialist will be unable to provide any reasons not to choose the best unless those reasons have already been given their due weight in determining which outcome in fact is the best. Allowing these reasons to count again as a reason not to choose the best would be a form of double counting and hence not rational. The argument I have suggested allows a satisficer to avoid being pushed into this corner by resisting the idea that all reasons to act can be captured in determining the goodness of outcomes. But the strategy works only if nonconsequentialism is correct.

Notes

1. The example is employed by just about everyone in this literature, starting with Michael Slote in "Satisficing Consequentialism," *Proceedings of the Aristotelian Society* 58 supp. (1984): 139–63, at 142. Slote claims to find it in the literature about rational economic behavior from which the term 'satisficing' is appropriated.

2. These terms are from David Schmidtz, "Rationality within Reason," *Journal of Philosophy*, 89 (1992): 445–66 and roughly are meant to contrast concern with one parameter of one's life and concern for one's life as a whole.

3. Thus, although I agree with one of the two theses defended by Christine Swanton in "Satisficing and Virtue," *Journal of Philosophy* 90 (1993): 33–48, I disagree with her about the other. She thinks that both of the following are true: (A) An action that has good enough results may rationally be preferred to one judged to have better results. And (B) An action that is good enough may rationally be preferred to one judged to be better. If preference here is a way of representing what one should choose, I agree with her about A, but not B for the reasons I list in the text.

4. Michael Slote, *Beyond Optimizing*, Cambridge: Harvard University Press, 1989, 130.

5. Slote at one point rather bafflingly seems to deny this. "Even granting that it is *more rational* to treat the times of one's future equally, . . . it does not follow and one may deny that it is always less rational to give some preference to the near. . . ." Michael Slote, *Beyond Optimizing*, 126. Unless the "always" in the last clause is doing all of the work, I can't actually figure out what this could mean. I take it that if *a* is more rational than *b*, it follows that *b* is less rational than *a*. Perhaps the statement is not so much his considered position as part of a broad survey of positions that seem to be made available by a careful consideration of the options.

6. Consider a parallel with the debate over consequentialism, construed as the thesis that right actions always bring about the best consequences. Someone who denies this should not be saddled with denying that the right action is always most right, or even that the right action is always best from the point of view of rightness (if that makes any sense), and certainly not with claiming that the right action can be less right than some alternative equally available under the circumstances. My saying this may depend partly on the argument below, so it would be somewhat unfair to rule out this way of construing that debate before I make the argument. Still, it is fair to underline that this way of looking at the issue is not mandatory for nonconsequentialists to accept.

7. This is what I take to be the moral of Joe Mendola's criticisms of Gauthier's notion of constrained maximization in Joseph Mendola, "Gauthier's Morals by Agreement and Two Kinds of Rationality," *Ethics* 97 (1987): 765–74.

8. David Schmidtz makes just this sort of complaint against Michael Slote: "If all we have is an intuition that an act makes sense, but cannot say what the act makes sense *in terms of*, then we would be jumping to conclusions if we said that we were approving of the act as rational." Schmidtz, "Rationality within Reason," 456. And earlier Phillip Pettit complains, "The irrationality of the policy first appears in the fact that whereas he could give a reason for choosing A – it is in his view the better option – he can give no reason for choosing B." Philip Pettit, "Satisficing Consequentialism," *Proceedings of the Aristotelian Society* 58 supp. (1984): 165–76 at 172.

9. None of these points are new with me. They are made, for example, by Schmidtz, "Rationality within Reason," 454.

10. Slote, *Beyond Optimizing*, 2.

11. Slote, *Beyond Optimizing*, 2.

12. Schmidtz makes all of these points. Schmidtz, "Rationality within Reason," 456.

13. Swanton, "Satisficing and Virtue," 37.

14. Swanton is careful to respond to this objection (raised to her by Pettit) that it would be rational to make the sacrifice for the friend even if there were no such weakening effects. Swanton, "Satisficing and Virtue," 38.

15. It appears to me that Swanton does not really address this sort of objection head on, although some of her remarks about the relevance of agent-centered prerogatives and restrictions on promoting value (on page 39, for example) are friendly to the sort of view I will suggest is needed to rationalize mere satisficing. I suppose where we differ is in this: She thinks that all that needs to be shown is that such restrictions make sense, and that virtue theory can be used to show that they do make sense. I think you need to do more than this: You must also show that the best way of making sense of nonconsequentialist prerogatives and restrictions on maximizing the good does not require folding all considerations including these restrictions into one agent-relative ordering of outcomes from better to worse. It is to this argument that the next several sections of this paper are devoted.

16. At least I believe that it is G. E. M. Anscombe who coined the term in the late 1950s, although her use of the term is not exactly the one that has gained currency. She seems to use it for any view that allows the goodness or badness or consequences to override duties that she wants to view as absolute side-constraints on actions. Thus a theorist like Ross, whom most of us would characterize as a nonconsequentialist, would be labeled a consequentialist if current usage abided by Anscombe's original implicit definition. But though she coined the term, subsequent usage has restricted it to theories that make the rightness of actions solely a function of the value of the action's consequences, and most of us limit the relevant sort of value to agent- and time-neutral value. G. E. M. Anscombe, "Modern Moral Philosophy," *Philosophy* 33 (1958): 1–19.

17. See, for example, David Sosa, "Consequences of Consequentialism," *Mind* 102 (1993): 101–22. Although Sosa is by no means the first to argue in this way, he is among the most liberal in the resources he claims on behalf of consequentialism insofar as he wants to count theories that include agent-relative features of outcomes in the specifications of the consequences of actions. Other theorists who preceded Sosa in advocating the use of agent-relative values in determining the rightness of actions have generally eschewed the label "consequentialist" for their theories, while recognizing that their theories are motivated by extensions of consequentialist rationales. For example, Amartya Sen calls his theory 'consequence based' because it allows the use of agent-relative values. Amartya Sen, "Rights and Agency," *Philosophy and Public*

Affairs 11 (1985): 3–39. John Broome revives the term 'teleological' for theories like his that allow agent- and time-relative orderings of outcomes to determine rightness. John Broome, *Weighing Goods*, Cambridge: Blackwell, 1991: 6. It should become plain from the text that I prefer to speak as Sen and Broome do rather than to continue to call theories that allow such relativized values 'consequentialist'.

18. Samuel Scheffler groups the first two sorts of reasons together as agent-centered restrictions on action and refers to the latter as agent-centered permissions. Samuel Scheffler, *The Rejection of Consequentialism*, Oxford University Press, 1982: 22–3. Nagel labels these three sorts of reasons respectively as "reasons of obligation," "deontological reasons," and "reasons of autonomy." Thomas Nagel, *The View from Nowhere*, Oxford University Press, 1986, 165.

19. Certainly the most influential argument of this sort is Bernard Williams's contribution to Smart and Williams, *Utilitarianism: For and Against*, Cambridge University Press, 1973.

20. The response that it isn't what consequentialism requires because two children saved as opposed to one will lead to population problems down the road won't help, because that strategy would also suggest that the parent should really not have saved anybody at all.

21. Sosa, "Consequences of Consequentialism," 115.

22. To see that this involves genuine agent-relativity, we need only notice that another parent may permissibly save another child who is his own rather than this woman's child but would not be required to prevent this woman from saving his child if he found her doing it.

23. Robert Nozick, who introduces the term 'side-constraint', notices in the same place that one could model side-constraints within a quasi-consequentialist framework if one allowed agent-relative betterness relations. See Robert Nozick, *Anarchy, State and Utopia*, New York: Basic Books, 1974, 29. John Broome has a nice discussion of how agent- and time-relative orderings are needed if one is to capture the workings of side-constraints in a teleological theory on pages 3–16 of *Weighing Goods*.

24. Our unwillingness to redirect the parent might merely reflect a judgment that we should not play favorites in such a situation. The point remains that the fact that one saving is the saving of a son does not commit us to viewing the state of affairs that includes it as any more valuable than any other alternative in which a life is saved, even when we think the father should save the son.

25. Frank Jackson, "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection," *Ethics* 101 (1991): 461–82, at 466–7.

26. James Dreier argues for the opposing view in part by suggesting that once we see that what I am calling "nonconsequentialist" theories can be captured with agent-relative orderings, we can defend them against the charge that they ignore the good, and also because we will be less apt to conflate objectivity with agent-neutrality, thereby making it easier for nonconsequentialists to be drawn into ways of thinking that give agent-neutral theories an upper hand. James Dreier, "The Structure of Normative Theories," *The Monist* 76 (1993): 22–40.

27. Michael Byron makes something like this point against people like Swanton who use nonconsequentialist moral theory to defend satisficing: "If, on the contrary, nonconsequentialism is true, then Swanton's conception of 'satisficing' is superfluous, since the truth of nonconsequentialism entails that choosing a suboptimific alternative is sometimes rationally permissible." Michael Byron, "Satisficing and Optimality," *Ethics* 109 (1998): 67–93 at 90–1. The paragraph that follows provides my answer to this objection.

28. Nagel, *View from Nowhere*, 176, and Michael Walzer, Chapter 16 of *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, New York: Basic Books, 1977, 251–68.

29. Thus, to contrast my paper with one of the others in this volume, my disagreement with James Dreier is not over whether one can construct the preference ordering that allows us to construe all choices as maximizing utility. I agree that we can. But I doubt that this is the notion of utility that someone who defends rational satisficing should use. I claim that an advocate of satisficing might use the notion of agent-neutral goodness from plausible consequentialist theories, and that if nonconsequentialism is right, it will make sense to satisfice with respect to such goodness. I think this is where Dreier would disagree with me.

30. Michael Slote, *Common-Sense Morality and Consequentialism*, London: Routledge, 1985, 3 and 136.