**MAXIMIZING, SATISFICING AND THE NORMATIVE DISTINCTION
BETWEEN MEANS AND ENDS**
Presented April 8, 2005

Robert Bass
Assistant Professor
Department of Philosophy
Coastal Carolina University

Decision theory, understood as providing a normative account of rationality in action, is

often thought to be an adequate formalization of instrumental reasoning. As a model, there is

much to be said for it. However, if decision theory is to adequately account for correct

instrumental reasoning, then the axiomatic conditions by which it links preference to action must

be normative for choice. That is, a choice must be *rationally defective* unless it proceeds from a

preference set that satisfies the axiomatic conditions.

The crucial feature of standard decision theory for present purposes is that it conceives

rational action as maximizing, as doing the best that one can in terms of satisfying one's

preferences. This provides poor guidance, because it mistakes a feature which may be present in

instrumental reasoning for a requirement upon rational choice. For maximizing to be possible,

the preference set in question must completely order a person's options. But that condition is not

met by actual preference sets, so maximizing is not always available. Given that it is not,

satisficing deserves attention, both as the most important alternative to maximizing and for the

lessons it can yield with respect to goal-directed action. With those lessons in hand, standard,

maximizing decision theory is reconsidered, with the aim of showing that it does not adequately

represent the normative distinction between means and ends.

How Incompleteness Undermines Maximizing

Some of decision theory's axiomatic conditions are uncontroversial; others are much less so. For present purposes, the most important condition is Completeness, the requirement that an agent's preferences completely order her options. This condition is needed if an agent is to *maximize* – she cannot be sure of selecting the best or most preferred option if her preferences do not order all her options. When Completeness is applied to elements in a preference set, however, the number of comparisons required, if each must be made explicitly, quickly becomes unmanageably large. If, in order to satisfy Completeness, one must explicitly compare each element in one's preference set to each other, then, for three elements, *A*, *B* and *C*, one needs to perform three comparisons: *A* to *B*, *A* to *C* and *B* to *C*. For four elements, six comparisons are needed, for five elements, ten comparisons, and so on. Evidently, unless the number of elements is small, this is going to quickly get out of hand. For example, for 50 elements, 1225 comparisons would be needed. Or, in general, for *n* elements, the number of comparisons needed is equal to the sum of the integers from zero to *n* - 1. If Completeness is extended to all the options that can be constructed under conditions of risk or uncertainty, a complete ordering would involve infinitely many pair-wise rankings. Since the agent cannot have explicitly performed all of these rankings, then, for her preferences to completely order her options, they must have an underlying structure which suffices to determine all the needed preferential relations.

But if explicit comparisons cannot do the job, it is hard to understand how an underlying structure could. The problem is especially pressing when novel elements must be integrated into

a preference set. This may happen in many ways, but what is important is the fact that novel experience introduces an agent to something which she did not previously rank preferentially at all. Once the agent is introduced to the new element, *C*, she will rank it in some way, say, against other elements, *A* and *B*. When she ranks *C* as being worse than *A* and better than *B*, she will have to get the relation *exactly* right for all possible gambles between *A* and *B*, as well as for all possible gambles in which *C* is compared to any other elements of the preference set – elements which may not have been considered at all in the initial ranking of *C*. Otherwise, the ranking of *C* will introduce intransitivities to the preference set, and thereby a violation of Transitivity – which is surely the least controversial of the axiomatic conditions. In essence, this is an argument from finitude. Getting the preferential relations right for each new element introduced into a preference set requires unlimited precision. Since it is not reasonable to believe in the consistent achievement of unlimited precision, it is not reasonable to believe that preference sets are complete.

Further, if a person's preferences do not completely order his options, there is little he can do to rectify matters. There are no obvious steps to take that would result in his coming to have a complete preference-ordering. The reasons for thinking his preferences are not complete are also reasons for thinking he cannot bring it about that they will, in the future, be complete.

Beyond that, even if there were steps a person could take to impose Completeness on his preferences, it is not clear that he would have any reason to do so, for the supposed reason would either depend upon a complete ordering of his preferences or not. If the former, then the argument for imposition is fatally compromised, while, if it is the latter, the incompleteness of the preferences leaves open the possibility that the reason will be undefeated, untied, but still not rationally decisive.

Since a complete ordering of all of a person's options is not inscribed in his preferences, anyone can find himself in situations in which there is no answer as to which option best serves his preferences. For maximization to be a general requirement upon rational choice, however, it must be possible to apply it to any decision problem that can be constructed from the elements of a person's preference set. Regardless of the options with which the agent is faced, it must be possible (in principle) to identify one of them as being at least as good as any other.

Since the agent's preferences do not fully order his options, maximizing cannot be appropriate to all choices, since it is not always well-defined what a maximizing choice would be. Further, the larger the scope of a choice – that is, the greater the extent to which it can be expected to have substantial and lasting effects – the more likely it is that maximizing will not be available. Choice of a career or of a mate provide good examples, for in each case other decisions will in turn depend upon the earlier decision.

Part of the point is that the further effects cannot be foreseen in detail and may therefore impinge in unforeseeable ways upon matters made relevant by one's other preferences. But so far, that is only a problem of uncertainty. There is an additional dimension due to the fact that one of the features of long-term plans is that their execution makes a significant difference to what the person is doing over the term of the plan and that *the person herself* is altered in the process. She engages in different activities, spends time with different associates, and acquires different preferences as a result of executing the plan. Importantly, some preferences relevant to the choice to adopt and execute the plan may be preferences the person *does not have* when the plan is adopted. The uncertainty runs deep: not only is the agent uncertain what the future may bring, she is also uncertain how the unknown future will matter when the time comes. The larger the scope of a choice, the larger is the set of preferences that may be relevant, and the set of

relevant preferences[1] probably no more than intersects with the *complete* set of the chooser's preferences at the time of choice.[2]

Satisficing and the Adoption of Means to Ends

If, then, maximizing cannot be applied to all choices – and, ironically, is least likely to apply where one would most like some clear-cut decision procedure – what can be done instead? The most plausible answer is that one should settle for *satisficing*.[3] The core idea is that the agent should seek and select an option that is *good enough*, rather than one that maximizes. It applies most naturally to cases in which an agent is still seeking an acceptable option. (Why select an option known to be worse than some other, simply because it is still good enough?[4]) Then a satisficer, rather than trying to determine what option is best in terms of all her preferences together, delimits some range within which a decision problem arises – such as what to have for dinner, whether to accept a job offer, whether to buy a house or keep searching – and then settles upon criteria such that, if they are satisfied, an option would, by her lights, count as *good enough*. Then, options are compared in light of these criteria, and the first to qualify as

---

[1] Set aside, for the moment, any concerns about how to determine in practice the membership of the set of relevant preferences. Then suppose that each of a pair of options would have different effects upon the preferences of the chooser such that, if one option is selected, the chooser will come to prefer *A* to *B*, whereas, if the other is selected, she will come to prefer *B* to *A*. Does it make sense to say that one of those preferences, to the exclusion of the other, belongs in the complete set? Surely, both are in some sense relevant and both have the *same* claim to be included, but if both *are* included, the preference set will not be consistent.

[2] Will the chooser have preferences about the ways in which her preferences are subject to modification in consequence of some far-reaching choice (which preferences can then feed back to provide additional criteria or desiderata for the choice)? Quite possibly, but there is no more reason to expect that these preferences will completely order her options than that her other preferences will do so.

[3] See Nozick (1981, 300), who cites Simon's 1957 *Models of Man*. The idea has been much discussed, both by Simon and others. See, e.g., Schmidtz 1995, Simon 1996/1969, and, without using the term, his 1990/1983.

[4] In special cases, it might make sense – for example, if there is neither a best member nor any tied for best in the set of available options. See Schmidtz 1995, 42-43.

good enough is selected.[5]

There are, of course, indeterminacies, intransitivities and practical dilemmas to which a satisficer is prey. Her choice in favor of one option and against others may be shaped by the order in which questions are asked and considerations brought to bear rather than by the relative merits of the options. It would be good to avoid such difficulties, if possible. *In principle*, the maximizer escapes them, but even at its best, the escape amounts to less than may appear. For a maximizer choosing in the face of risk or uncertainty, the maximizing choice may be to select the best member of a limited set of options, consisting of, say, *A* and *B*. It may still be true that, had he considered a third option, *C*, he would have ranked it above both *A* and *B*. Being a maximizer does not protect an agent against the possibility that actual decisions may depend upon the order in which options are presented or upon other extraneous factors, rather than upon the relative merits of the options. More importantly, the promised escape from practical dilemmas is only an illusion in any case since maximization is not always possible.

A formal analysis of satisficing[6] need not be pursued here, except to note the interesting point that it appears that any rationale for satisficing must itself be a satisficing rationale. No argument can be mounted that satisficing is the best that can be done (given uncertainty and incomplete preference orderings), for, apart from the fact that 'best' may have no determinate reference in the face of incompleteness, its success would be its failure. If there *were* a sound

---

[5] If the criteria turn out to appear *too* easy to satisfy, they may be revised upward, or if too difficult, then downward. In either case, what is "too difficult" or "too easy" is itself at least implicitly a function of a satisficing judgment – that the effort and resources devoted to the search is or is not good enough. See Nozick 1981, 300.

[6] See Schmidtz 1995, Chapter 2 and especially 55-57. Relying upon the fact that satisficing applies most naturally to the search for an acceptable option, Schmidtz shows that satisficing can be distinguished from maximizing over the issue of when to terminate the search. The maximizer will continue to search as long as the expected utility of searching is greater than the expected utility of the best option found up to that point. The satisficer, relying upon already-established criteria, will terminate the search as soon as some option that is good enough is found.

general argument that satisficing is the best policy, that would assimilate satisficing to maximizing. Satisficing would then be what maximizing under those conditions amounted to. Satisficing can only be a genuine alternative if its rationale is something other than that it is the best procedure for selection among options.[7] And if the rationale is not that it is good enough, or satisficing, what could it be?[8]

Much can also be done in the way of providing a rationale for satisficing by exhibiting problems with the maximizing model, but that is sufficiently complete for present purposes. What is important is to attend to a feature of satisficing that suggests the deepest problem with standard, maximizing decision theory.

How does a satisficing agent guide her action? Within some domain, she selects as an objective some state of affairs which she believes can be brought about or promoted through her action. She is guided by her judgment that the selected state of affairs is good enough, that it answers satisfactorily to her desires and preferences. In other words, she selects a goal and, barring alteration of the goal itself, guides subsequent action in that domain by its suitability for the promotion of that goal rather than by its suitability for maximizing her preference-satisfaction in general.

Thus, there are two distinguishable stages in the deliberation by which a satisficer guides

---

[7] Schmidtz says that satisficing can only be of instrumental value "because to satisfice is to give up the possibility of a preferable outcome, and giving this up has to be explained in terms of the strategic reasons one has for giving it up." (1995, 45) Though he makes it clear that he thinks that the strategic reasons for (sometimes) satisficing are rooted in maximizing from a larger perspective, the conflict with the view that incompleteness may make a maximization requirement inapplicable is more apparent than real, since he admits (46) that there may often be no optimum from a global perspective: an agent may have to make a choice when nothing unequivocally favors one option over another.

[8] The various axiomatized methods for choice under uncertainty do not provide alternative, non-satisficing routes to the selection of satisficing because they are all ways of identifying some maximand which completely orders options. The satisficer does not have any general procedure for inducing a complete ordering over options. (It is an interesting question for further exploration whether the selection of one of those methods for choice under uncertainty might presuppose satisficing in that there is no proof that one of those methods is best.)

her action. First, there is goal-selection carried out in light of the agent's preferences, but it is not assumed to be necessary either that selection of the particular goal or even of some goal or other (then and there) is a maximizing choice. The fact that the goal is selected as being good enough, as answering well enough to her preferences (which will not normally fully order her options), has the important implication that it need not be abandoned instantly should something better or apparently better come along. Since it was not selected for being the best, even proof that it is not the best need not lead to its abandonment. The deliberation relevant to abandoning a goal in favor of something else will have a *dissatisficing* structure. It will be appropriate to abandon a goal when it turns out to be *bad enough*. For a satisficer, there will be a gap between barely finding something else to be better than a currently pursued goal and appropriately abandoning its pursuit.[9]

Second, once a goal has been selected, action within the relevant domain is guided by its relation to that goal rather than by maximization. An agent who has embarked upon an investment plan may have chosen to set aside a given percentage of her income every month. Having decided that, she does not reconsider what to do with that portion of her income whenever an unanticipated opportunity for expenditure arises. She does not, in a typical case, ask whether she would really be better satisfied, all things considered, with new furniture.

There is a clear sense in which the satisficer selects a goal and then guides her action by its relation to that goal. The issue is whether the *maximizer* can similarly distinguish his goals from the means appropriate to them – more precisely, whether he can distinguish the differing

---

[9] There are interesting comparisons to be made with Joseph Raz's conception of authoritative reasons as pre-emptive: "the fact that an authority requires performance of an action is a reason for its performance which is not to be added to all other relevant reasons when assessing what to do, but should exclude and take the place of some of them." (Raz 1986, 46; emphasis in original omitted) Adopting a goal is analogous to recognizing an authoritative reason and pre-empts other reasons that would have been relevant had the goal not been adopted.

ways in which preferences with respect to outcomes and preferences with respect to the steps involved in bringing about those outcomes are relevant to his choices. If the distinction cannot be adequately drawn, then there will be a sense in which the maximizer *cannot* be said to guide his actions in terms of his goals.

The Maximizer's Problem with Plan-Execution

Consider the following all-too-common problem. An agent has adopted a plan at one time to bring about a preferred outcome at a later time. Execution of the plan requires performance of a particular action (the Step) at an intermediate time. At the intermediate time, there has been no change in relevant information available to the agent nor any unforeseen change in the agent's preferences, but as the time approaches, the agent strictly prefers not to take the necessary Step. In addition, the preference change with respect to the Step was itself foreseen when the plan was adopted.

Such situations are familiar. An example might be deciding upon a diet. There is an envisioned outcome, losing weight, ranked above other future outcomes and a necessary step, refraining from between-meal snacks. In addition, when the plan is adopted, it is recognized that there will be temptations to snack between meals: when the Step must be taken, the agent will prefer snacking to sticking to the diet. On one hand, it appears that the agent's reasons for taking the Step are just the same as for initially adopting the plan – no unanticipated information or preference has entered the picture. If the plan was initially well-conceived, the agent ought to take the Step.

On the other hand, now that the prospect of snacking is immediate, the agent does *not* prefer to lose weight. He would, right then, rather snack than lose weight. Why must he be

bound by his preferences of a few hours earlier? If it is rational for him to guide his actions by his preferences, why are the earlier preferences decisive, while those actually experienced when the choice must be made are discounted?[10]

Most people – however difficult they find it to carry through in practice – suppose that the former argument is better: Having made a reasonable plan, and in the absence of relevant additional information not already taken into account in the formulation of that plan, it is reasonable for a person to take the necessary steps to implement the plan, even if those necessary steps are dispreferred at the time they must be taken.

However, according to standard decision theory, this misdescribes the situation. It is not that the first argument is invalid, but that it depends upon a false premise. In standard decision theory, the only reasons people have for action are based on preferences and expected consequences at the time a choice is made. A decision cannot rationally depend – except insofar as this affects current preferences and expectations – upon a past event such as having adopted a plan. Thus, if the step needed to carry out the plan is such that one would prefer not to take it at the time of choice – when the step must be taken or not – then one has reason not to take the step. But if this was really foreseen when the plan was adopted, then it was not reasonable to adopt the plan since its execution depends upon the taking of an unreasonable step. One who accepts the rationality-defining postulates of standard decision theory should either not have formulated a plan aiming at that goal or else should have made provision that every step would be preferred to its alternatives when it would have to be taken. For standard decision theory, reasonable plans

---

[10] Does he still *have* the preference to take the Step, even if it is not motivationally salient? Perhaps it should be allowed that he may, but if so there is at least an apparent conflict among his preferences, and it is not clear which should govern his choice.

contain only feasible steps, where feasible steps are all preferred, when they must be taken, to

their alternatives. If that requirement is not met, then the plan was not reasonable in the first

place.[11]

This seems unsatisfactory. If standard decision theory is correct about such situations,

there may be an outcome which an agent would like to achieve, a plan that, if executed, would

achieve that outcome, and it may be that if the plan were executed and the outcome achieved, the

agent would be glad she had adopted the plan and taken all the necessary steps, but nonetheless,

the agent cannot rationally adopt the plan because it incorporates infeasible steps. Her best

available options are either to give up seeking that outcome or to undertake special arrangements

to make sure all the steps are feasible. Either option represents a cost, whether in the form of

giving up the chance to obtain her most preferred outcome or in the form of making special

provisions to avoid having to take infeasible steps.[12]

Why does this problem seem so difficult for the maximizer of standard decision theory?

There are two points to note before trying to answer. First, the question is not just why most

people sometimes find it hard to carry out their plans. Imperfect rationality adequately accounts

for that. Rather, the question is about why *ideally rational* agents would find themselves

apparently having to settle for second-best.[13] And second, it is specifically a problem for rational

agents *as conceived by standard decision theory*. If, as has been argued, the conception of

rationality embodied in standard decision theory is not normative for human beings, it may be

---

[11] See McClennen 1990, especially chapters 12 and 13.

[12] A further concern is that the feasibility-insuring provisions might themselves be so costly that, if they are necessary to achieve the outcome, then the outcome is not worth achieving.

[13] Note that the problem is *not* one of weakness of will. It is not that one cannot bring oneself to do what one knows is best or cannot resist temptation to do what is worse. It is that what is rationally required seems not to be best.

possible to address the problems associated with taking steps to achieve a goal in ways not open to the maximizer.[14]

The Normative Distinction Between Means and Ends

The reason the problem seems difficult is that standard decision theory has no satisfactory way of making the normative distinction between ends and means. If the distinction could be made, there would be conceptual room to hold that ends provide reasons for adjusting means but not *vice versa*. To see what the problem is, consider where or how such a normative distinction might be represented. There are two plausible candidates, that ends are to be characterized in terms of outcomes of actions or in terms of intrinsic preferences.

Suppose ends are identified with outcomes and hence means are identified with steps that contribute to bringing about those outcomes. Then, the problem is simple: Though an answer can be given as to which steps contribute to what outcomes, all normativity vanishes because the fact that a step contributes to an outcome will not provide any reason for taking that step or for avoiding its alternatives. Any step and any combination of steps will lead to some outcome or other. What is needed, at minimum, is some way of discriminating *among* outcomes, to identify one or some as ends, rather than others, and therefore to enable the identification of some options, rather than others, as means to those ends.[15] In short, in addition to the identification of outcomes and contributory steps, something else is needed to represent the normative force of

---

[14] Satisficers do not face the same problem, at least not in so acute a form, for they are not automatically subject to criticism for making non-maximizing choices and therefore not for taking counter-preferential steps. (Being a satisficer may not be sufficient to deal with the problem in all its forms.)

[15] And that is just the beginning, for many features of outcomes of action are not intuitively part of any end pursued in a given course of action. Typing rearranges small particles on the keyboard, but the arrangement or rearrangement is not what the typist is aiming at in typing.

ends.

It might be thought that the other factor can be readily supplied. Consider *wholly derived preferences*, or *derivative preferences* for short. At a given street corner, a person prefers turning left over turning right because she prefers one grocery store to another. If not for the preference between stores, she would have no preference for turning one way over the other. On pain of infinite regress, however, not all preferences can be wholly derived; some must be non-derivative or *intrinsic preferences*.[16] The proposal, then, would be that ends are to be characterized in terms of intrinsic preferences. Ends will be intrinsically preferred to their alternatives, so, once ends are securely identified, contributory means can be considered. It can be shown what derivative preferences an agent should have and act upon in light of her intrinsic preferences.

The problem with this is that no role is left for temptation. Return to the story of the diet, and consider the readily generalizable case of George, whose end or goal is to lose weight. So far as he has only derivative preferences between steps or means, the only explanation for his not taking a step that is better than available alternatives at contributing to his ends must be in terms of misinformation, ignorance, or inadvertence. He will certainly have no motivation to take a step that either leads away from or less effectively toward his ends. But then, whence is the temptation to snack? Surely, yielding to temptation is not a matter of an accidental misstep on the way to his goals.

The answer must be that George's preferences with respect to the steps to be taken are not

---

[16] There may be single preferences, where *A* is preferred to *B*, which are wholly non-derivative in the sense that only the preference for *A* over *B* is relevant to any choice between the two. But it may be that a preference is not wholly derivative without being wholly non-derivative. The preference relation between the two may be part of some set of mutually supporting or interlocking preferences such that *A* would be preferred to *B* if nothing else were at stake, but that if something else were at stake, the preferential relations could be altered. Since no important part of the argument turns upon the distinction between wholly and partially non-derivative preferences, 'intrinsic preference' will be indifferently employed to cover both.

wholly derived. He is motivated to snack rather than stick to the diet because he has some intrinsic preference for snacking, then and there. If so, there are two possibilities, that the preference for snacking can or else that it cannot be integrated into a consistent ordering with George's other intrinsic preferences. If it cannot, then there is no consistent set of intrinsic preferences to identify as the relevant end or ends, and therefore none in terms of which to regiment means.

Matters are no better, however, if it is supposed that George's preference for snacking can be integrated into a consistent ordering. For then, at least *prima facie*, the act in question is not, strictly speaking, one of *yielding* to temptation; rather, it is an act licensed by its service to his ends. There is an intrinsic, rather than merely a derived, preference for what is being called "yielding" and, since ends are to be identified in terms of intrinsic preferences, no genuine yielding after all.

Now, it might be supposed that room for the possibility of yielding to temptation can be found in the thought that the snacking to which George is tempted is contrary to what he really or most prefers. Though his intrinsic preferences, including the preference for snacking, can be integrated into a consistent ordering, snacking then and there does not serve them.

This is very puzzling. Though there are several possibilities here, none seems adequate. To begin, what is the other element of George's preference for snacking: what is snacking preferred *to*? Presumably, it is preferred to not snacking. Also, however, sticking to the diet, which requires not snacking, is preferred to snacking. If that is not simply to amount to an inconsistent set of preferences and therefore to an inconsistent set of ends identified in terms of those preferences, there must be some sense in which the preference for sticking to the diet is what sets George's end while the preference for snacking does not. Since both preferences are

intrinsic, they cannot be distinguished on that basis.

The solution must be that adhering to the diet is what George most prefers, but how is that to be understood? The meaning is not that the diet-adherence preference has greater introspectible intensity, for in those terms snacking may well be what George most prefers. Nor will it do to accept, without further elaboration, the formulation that the act of snacking is contrary to what George most prefers, because, for the decision theorist, there are only formal limits on what may enter into a utility function, and, subject to those constraints, any preference is to be considered on the same terms as any other. The set of his preferences is equally consistent if he snacks and alters the preference for adhering to the diet as if he refrains from snacking in order to adhere to the diet.

Two attempts to provide the needed elaboration on what can be understood by what George most prefers, are that the answer is to be found through appeal to higher-order preferences, or that utility-maximization may be served by adherence to a rule.[17]

Consider the first. Suppose George not only has preferences but preferences over his preferences – that, along Frankfurtian lines, he has preferences about which preferences will be effective in guiding his action (Frankfurt 1988). He may then have the intrinsic preference for sticking to the diet and also the intrinsic preference for snacking, but what makes the difference between them is that there is a higher-order preference, presumably also intrinsic, that the diet-adherence preference be effective in action.

Consider the second attempt. In the spirit of something like rule-consequentialism, it may be said that better consequences can be had by adhering to a rule than by calculating the best

---

[17] They were suggested by Malcolm Murray and Jeremy Koons, respectively.

action on a case-by-case basis. The proposal would be that George acquires the goal of dieting

when he adopts a diet-adherence rule. Then, though there is an intrinsic preference for snacking,

what makes it the case that dieting is his end rather than snacking is that the snacking is contrary

to the rule.

Does either of these solve the problem? It appears not. In both cases, the problem

remains of comparing the importance of items intrinsically preferred – in one case, snacking

versus the effectiveness of the diet-adherence preference or, in the other, snacking versus

adherence to the rule. Why is it the rule or the higher-order preference that sets the end? The

problem is the same as at the beginning, with what, in each case, appears to be a conflict between

the members of a set of intrinsic preferences. And there are the same two possibilities. George's

intrinsic preferences with respect to dieting and snacking either can or cannot be integrated into a

consistent set. If they cannot, there is no end to be defined in terms of his preferences. If they

can, there is still no reason that has been given for identifying his end with sticking to the diet

rather than with snacking.

What is needed is some further explication of the sense in which the preference for

sticking to the diet is supposed to be of greater weight or importance than the preference for

snacking. That explication has not been, nor can it be expected to be, forthcoming. More

precisely, it will not be forthcoming in terms that can be represented within standard decision

theory, for the explanation being sought is one of the normative distinction between means and

ends, not of the psychology or phenomenology of preference or desire. When it is asked why

adherence to the diet is more important than snacking, the question is why George *should* abstain

from snacking, and the answer to that lies in the fact that losing weight is George's goal or end.

It is because losing weight is the goal that adherence to the diet is more important than snacking,

not because adherence is more important that loss of weight is the goal. There will be no answer in terms of preferences alone.[18]

And that is the deepest problem with standard decision theory. Whether explicitly or not, the theory seeks to be reductive about ends, to account for them in terms of the satisfaction of preferences and the like.[19] But the attempt to understand rational choice only in terms of maximizing the satisfaction of preferences ultimately leaves the theory unable to express or represent the normative distinction between means and ends. Taking instrumental reasoning seriously requires going beyond decision theory.[20]

---

[18] Nor will there be an answer in terms of (just) preferences combined with beliefs and expectations. Preferences, beliefs and expectations, of course, enter into the selection of goals, but the role of goals or ends in the guidance of choice is not captured in those terms alone.

[19] A non-reductive account of ends might be along the lines of Bratman's planning theory of intention. Though he does not typically speak in these terms, roughly, an objective or goal is what intentional action is guided toward, and an intention is "a distinctive attitude, not to be conflated with or reduced to ordinary desires and beliefs" (Bratman 1999, 10) – nor, it might be added, should it be conflated with or reduced to preferences.

[20] Acknowledgements.

References

Bratman, Michael E. 1999/1987. *Intention, Plans, and Practical Reason*. In *David Hume Series of Philosophy and Cognitive Science*. Stanford: Center for the Study of Language and Information.

Frankfurt, Harry G. 1988. Freedom of the will and the concept of a person. In *The Importance of What We Care About*. Cambridge: Cambridge University Press.

McClennen, Edward F. 1990. *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge: Cambridge University Press.

Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge: Belknap Press of the Harvard University Press.

Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.

Schmidtz, David. 1995. *Rational Choice and Moral Agency*. Princeton, N.J.: Princeton University Press.

Simon, Herbert A. 1990/1983. Alternative visions of rationality. In *Rationality in Action: Contemporary Approaches*, edited by P. K. Moser. Cambridge: Cambridge University Press.

_____. 1996/1969 [Third edition]. *The Sciences of the Artificial*. Cambridge: The MIT Press.