SELF-INTEREST, AUTONOMY, AND THE PRESUPPOSITIONS OF DECISION THEORY

Mariam Thalos

The foundations upon which many contemporary economists, social scientists, and contractarian political theorists build is that of rational choice theory, known also as decision theory.¹ This discipline endeavors to answer the oldest of arguments raised against the very possibility of a social science: that there can be no genuine science of individuals with free will, since the voluntary actions of such beings cannot be covered by causal laws. Decision theorists, accepting the premise of this argument, appeal instead to noncausal laws predicated on principles of success-oriented action, and use these laws to produce substantive and testable predictions about large-scale human behavior. The primary directive of success-oriented action is maximization of some valuable quantity.² Many economists and social scientists use the principles of decision theory to explain social and economic phenomena, while many political philosophers use them to make recommendations on questions of public policy.

Critics of decision theory, many of them feminists, complain that principles of decision theory are, nevertheless, an inappropriate foundation for social science and political theory because they rest on illegitimate assumptions. These illegitimate assumptions, according to the critics, include the following two propositions: (1) that rational decision makers are purely self-interested; and (2) that rational decision makers are autonomous entities — answering only to their own beliefs and desires. I shall call (1) the Self-Interested Deliberator assumption, and (2) the Autonomous Deliberator assumption.

Rationality, in the most general and uncontroversial sense, is the capacity to consider reasons for and against. And prudence (another name for the capacity for success-oriented action) is the capability for choosing suitable and efficient means in the service of independently acquired ends; prudence simply consists in conformity to some (more or less) effective means-ends calculus. What is the relationship between rationality and prudence? According to one very famous view, the terms “rationality” and “prudence” name the same thing. Bertrand Russell, for example, writes that “‘Reason’ signifies the choice of the right means to an end that you wish to achieve. It has nothing whatever to do with the choice of ends.”³ And Herbert Simon writes, “Reason is wholly instrumental. It cannot tell us where to go; at best it can tell us how to get there. It is a gun for hire that can be employed in the service of any goals we have, good or bad.”⁴ But according to another view, put forward most famously by Kant, prudence is but a special requirement of rationality,
while rationality as such makes further, independent, demands as well. In particular, it places restrictions on which ends may be self-chosen, and it commands pursuit of certain ends, such as duty and knowledge, whether they are self-chosen for self's sake or not. So a rational individual, on this contrasting view, has ends that go beyond, and perhaps even override, the self-chosen.

In this essay I shall use the term "rational" in the neutral sense according to which a rational individual gives consideration to reasons, and I shall have nothing to say about the demands of rationality as such. What I shall do instead is examine prudence, and show that its requirements are more complex and philosophically interesting than has so far been suspected. Specifically, I shall identify two independent formulations of the Autonomous Deliberator assumption, one that concerns choice of ends and another that concerns prudence. Then I shall argue that decision theory does not rest on either formulation of the assumption, any more than it rests on the Self-Interested Deliberator assumption. For that Autonomous Deliberator assumption which concerns prudence cannot be an unconditional requirement of a satisfactory theory which answers the question: What is the procedure for choosing the right means to independently acquired ends? Hence no theory of the rational individual which begins from the premises of decision theory, nor any social or political science which builds further upon such a theory — unless it undertakes an independent commitment to either the Self-Interested Deliberator or the Autonomous Deliberator assumption — rests on either. It shall therefore be my thesis that the framework of decision theory does not presuppose anything substantial concerning the aims or decision procedures of the rational individual. Decision theory is, instead, a framework within which competing models of rational human aims and procedures can be constructed, and (if all else goes well) subjected to empirical test.

I. Homo Economicus: Cheers and Troubles

Kant, as is well known, put forward the proposal that autonomy is an all-or-nothing property, predicated of the rational will on a priori grounds. And, in the same spirit, others have argued that it is self-contradictory to say of a nonautonomous will that it is guided by reasons, and so is rational, since a nonautonomous will does not direct itself at all but instead takes direction from another agency (someone or something recognized as in authority over it), and so has no use for reasons. Autonomy, according to this first family of views, would appear to denote a certain freedom from constraints vis-à-vis choice of ends.

Feminists, on the other hand, disagree among themselves over whether autonomy is in any respect desirable, and whether women in particular ought to achieve autonomy wherever this is possible. There are feminists who believe that the trouble with autonomy is that women do not enjoy enough of it in real life, others who take a dim view of autonomy on grounds that it is tainted with (unworthy) masculine ideals, and still others who believe that autonomy is all too often confused with (both unrealistic and undesirable) separation from other human beings. In the preponderance of feminist writings, however, autonomy is characteristically contrasted with positive conceptions of interpersonal relationships and community values. Autonomy, as conceived by feminists who oppose it, is a psychological property predicated of persons in various degrees, and serving both as a descriptive term and a normative ideal held up for
emulation, especially to males. Jean Baker Miller, for example, writes that autonomy "carries the implication — and for women therefore the threat — that one should be able to pay the price of giving up affiliations in order to become a separate and self-directed individual." And Christine de Stefano describes the "Western, liberal, individualist subject whose autonomy consists in her ability to disengage from social context and relationship." The psychological ideal of autonomy criticized by these feminists includes not only self-control and independence of judgment, but also an unwavering adherence to abstract principles, even when this conflicts with fellow feeling for others. Autonomous agents, according to this second family of views concerning autonomy, act from detachment. And this is the ideal of autonomy which is presented as the summit of moral achievement in the developmental psychology put forward by Lawrence Kohlberg and criticized by Carol Gilligan.

Gilligan herself, unlike many who cite her work, does not consider the alternative she sketches to Kohlberg’s picture of the morally mature individual as incompatible with it. She writes that the empirical cases she discusses "are intended to illustrate two cross-cutting perspectives that do not negate one another but focus attention on different dimensions of the situation, creating a sense of ambiguity around the question of what is the problem to be solved." Even so the Autonomous Deliberator has met with more contempt than admiration among feminists reviewing the principles of decision theory and the social and political theories built upon them. Many, following Virginia Held, refer scornfully to decision makers who fit the description of Self-Interested and Autonomous Deliberator as "economic men." Held, who conceives of "economic man" as the ideal rational contract seeker of contractarian theory, writes: "Rational choice theorists point out that their theories are formulated for just those situations where individuals do seek to maximize their own interest and are uninterested in each other’s interests. . . . But the questions I am trying to raise . . . have to do with how we ought to treat, conceptually, a great variety of human relations. Of course we can, theoretically, treat them as contractual, but should we do so? Is it plausible to do so?"

II. The Presuppositions of Decision Theory

Decision theory divides into two independent parts. The first is a theory of preference, the second a theory of choice. The standard (von Neumann) theory of preference is a collection of formal rather than substantive principles that concern coherent ordering of preferences vis-à-vis ways the world may be — what I shall refer to as prospects. These principles, first of all, prescribe division of prospects into classes. Among members of the same class, one is indifferent, but not between members of different classes; the classes are therefore known as indifference classes. The strength of preference between members of different classes is then assumed to be quantifiable. Quantification of preference intensity entails a potential to represent an individual’s preferences by what is known as a utility function. Utility is then assumed to be the quantity which rational individuals aim to maximize.

But no theory of preference, standing on its own, can make proclamations as to which course of action is advisable under circumstances in which an individual cannot, simply by holding or revealing the preference, achieve the most favored among incompatible prospects. In other words, the imperative to maximize is not specific enough to count as the theory of
prudence — a theory of "the choice of the right means to an end you wish to achieve." An independent theory of choice, based on evident principles which admit of being extended to situations in which strategy is called for, must be brought forward to complete the treatment of rational decision.

The Self-interested Delibrator assumption, to which we will give more attention presently, clearly belongs to the province of the theory of preference, but where does the Autonomous Delibrator assumption belong? If we understand autonomy as the property of being unconstrained by others in choice of ends, then clearly it belongs to the theory of preference. But why then should feminists find it objectionable? And why should they proclaim that it conflicts with communitarian values? For the aspiration to be unconstrained by others in one's choice of ends does not conflict with promotion of collective advantage; if anything, one might suppose that the state of being unconstrained precedes community-oriented enterprises, especially those entered into by self-directed individuals. I shall argue that there is a formulation of the Autonomous Delibrator assumption which does conflict with communitarian values, but that decision theory does not rest on it.

Just as there are two divisions of decision theory, so there are two formulations of the autonomy proposal. The first, and that which I have here attributed to Kant and his followers, concerns preference. I will call it autonomy of the first kind. It is the thesis that, for a rational individual R, there are no others whose preference for R's preferring prospect X to prospect Y, gives R a reason to prefer Y to X. The second formulation of autonomy, which I shall call autonomy of the second kind, concerns choice of action. It is the thesis that, for any rational individual R, there are no others whose preferences for R's performing action A (which might contribute to bringing about a favored prospect Z), gives R a reason to perform A. The question whether to embrace autonomy of the first kind is a question that belongs to the province of preference, and the question whether to embrace autonomy of the second kind belongs, at least in part, to the province of choice.

It is autonomy of the second kind which causes trouble for feminists. For those who find autonomy objectionable do so largely because they believe that it stands in the way of collective action for mutual benefit. Collective action is what the detached person is worst at achieving, because he cannot be trusted to carry out his part of a collective action. The detached person is not untrustworthy because he is self-interested; rather, he is untrustworthy because he cannot manage to coordinate with others for the sake of all. The ability to coordinate with others does not presuppose that one has no self-interested or individual aims, but only that one can aim with others in a collective fashion. That one can aim with others presupposes that one can put individual aims, whatever their content, to one side, at least instrumentally. And the ability to put aside one's aims does not require that one deny them. It presupposes only that one gives some weight to the fact that others have preferences for one's performing certain actions. It rejects, therefore, autonomy of the second kind. As should, I shall shortly argue, any adequate theory of choice.

The standard von Neumann theory of preference leaves it as a matter of individual taste where a particular prospect should fall in the preference scheme. Thus it passes no negative judgment on an individual who, for example, prefers the destruction of a distant civilization to the pricking of her finger. By the same token, it passes no negative judgment on
an individual who prefers the advancement of others, whether loved ones or strangers, to the advancement of self. For it passes no judgment of any kind on the content of preferences; it neither censures nor approves, but only concerns itself with coherence. Hence decision theory, contrary to the critics’ complaint, does not require that rational individuals be purely self-interested.

A weaker proposal than pure self-interest is “mutual unconcern,” which is satisfied when utility functions are independent of one another. One expression of this proposal is by John Rawls, who writes that people are to be “conceived as not taking an interest in one another’s interests.” But the proposal of mutual unconcern is not presupposed by decision theory itself, since (to reiterate) decision theory does not place substantive constraints on the content of preferences. When, in decision theory, it is assumed that the actions of a rational individual shall spring from their preferences, this is a formal constraint only, since it does not prohibit those preferences from being other-regarding. And this fact is well known to many who stand in the traditions targeted by feminist critics of decision theory: Kantians and utilitarians in particular acknowledge not only that it is rational (and altogether commendable) to have a vested interest in seeing others advanced but also that it is rational, albeit not always profitable, to act out of moral rather than self-interested reasons, especially when one can derive no personal satisfaction from seeing others advanced before oneself. And they acknowledge, furthermore, that to act out of moral reasons can be an autonomous act.

The fact that the presumption, in decision theory, that the justification (and hence the explanation) of a rational agent’s actions must lie entirely with her preferences is a purely formal rather than substantive one, is a decisive reason for dismissing the charge that decision theory presupposes the Self-Interested Deliberator assumption. And it is a decisive reason also for dismissing the charge that decision theory presupposes autonomy of the first kind. For, since decision theory makes no substantive demands of any kind vis-à-vis preferences, a fortiori it does not demand that preferences be self-given. In other words, since decision theory places no constraints on choice of ends, the constraint that choice of ends be unconstrained in a particular way is not part of decision theory.

What, now, of autonomy of the second kind? In order to determine whether decision theory presupposes this second kind of autonomy we shall have to examine presumptions made in the province of choice. However the theory of choice currently on offer is a mixed bag of tricks, each trick with a restricted set of application conditions. Moreover the ranges of application for a number of principles of choice continue to be heatedly disputed. Even so, I shall argue that autonomy of the second kind cannot be a presupposition of a satisfactory theory of choice. I shall do this by showing that decision theory ought to look with favor on deliberative strategies that, under appropriate circumstances, permit the deliberator to be subject to influence by the desires and expectations of other deliberators.

III. CONCERTING EXPECTATIONS

It is customary in decision theory to characterize decision opportunities and those on whose doors they knock in the most slender (preferrably quantifiable) terms. This is for the sake of isolating only the hard core of features that distinguish one dilemma from another, in order thereby to be in a position to treat like dilemmas alike. It is a point of policy, therefore, to define a dilemma entirely in terms of
(anticipated) premium structure, and decision makers simply as “ideal” principals, whose deliberations hover about only those matters that make their dilemmas what they are. The theory of rational choice thus turns attentions away from such features as might distinguish one real-life situation from another in which (anticipated) options and premiums are the same, and it bids decision makers do likewise. Thus while decision theory never denies that each individual and circumstance is unique in any number of ways — for example, in historical antecedents — nevertheless it requests our consent to the proposition that the distinguishing qualities of decision makers and their circumstances should not be allowed to affect rational decision making, on grounds that a rational solution to a dilemma ought to be supported by reasons and reasoning which any person of reason, in sufficiently similar circumstances, would find compelling.

I will refer to the premium structure of a dilemma as its abstract form or structure, and to true-life dilemmas as instances of such forms or structures. The customary presupposition made by many decision theorists is therefore the assumption that a solution to a dilemma is a solution to an abstract, repeatable structure, and that such a solution will be a solution to all instances of that structure. This is an assumption that will be challenged in this essay, much as it is challenged in Thomas Schelling’s pioneering work, The Strategy of Conflict;\textsuperscript{18} and with it the proposition that decision theory must assume autonomy of the second kind.

In many dilemmas, the configuration of premiums alone does not distinguish any combination of strategies uniquely. For example, suppose you and I, without communicating, are each given only one opportunity to name either “heads” or “tails;” we win a certain prize so long as our selections match. The outcome of “heads/heads” is therefore just as agreeable to us as the outcome of “tails/tails.” The trick is to find our way, intellectually but without explicitly communicating, toward concerting action. The problem, however, is that if in individual deliberations we can bring to attention only the premium configuration of our dilemma — its abstract form — we cannot make any choice at all, or will perforce resort to the flipping of coins (leaving matters to chance), since there is no basis in the premium structure for favoring one choice over another. But the heads-tails dilemma is in reality quite simple to resolve for almost any pair of players. Both will choose “heads,” will they not?\textsuperscript{19} But why does each choose “heads”? Because the choices are labeled differently, and because “heads” enjoys a psychologically more prominent position in the heads-tails scheme, which fact is itself well known; so each partner knows that “heads” will leap to the other’s mind before “tails.” And are they each not to be admired, from the point of view of prudence, for selecting “heads,” having recognized some — any — mutually recognizable reason to expect the other to expect oneself to expect the other to do likewise?

Thomas Schelling writes that any account worthy of the name of a theory of strategy must acknowledge that the heads-tails dilemma has a genuine solution in real life, a solution not found by looking at abstract form alone. It is, after all, real life dilemma-instances that theories of strategy must resolve. It does no good to have a theory of strategy whose solutions to dilemma forms either do not apply to, or else give incorrect resolutions of, the true-life instances of those forms. Thus any respectable theory must recognize that certain solutions to problems of strategy are solutions precisely because they enjoy
"a signalling power, a means of tacit communication, that is available to the two players [undergoing deliberations in isolation] to facilitate their tacit cooperation when failure to coordinate choices would be serious."  

In some instances a solution is the solution simply because it presents itself, via its prominence, as such. Its prominence — for whatever reasons it might happen in that instance to be prominent — serves as a clue to the players. Symmetry, for example, might focus the expectation of two players, as it usually does in the following dilemma: two players are to name a sum from $0 to $100; if the sums named add to $100 each receives the sum claimed, otherwise nothing. But if symmetry (a 50-50 split) is the solution in this problem, symmetry is the solution precisely because symmetrical play is anticipated by the players; it is not true that symmetrical play is anticipated because symmetry is the correct solution. We could not fault a pair of players, each of whom independently arrives at a 25-75 division favoring a certain one of them. On what basis could an uneven division, agreed to tacitly by a certain pair of players, be faulted?

And symmetry is not the only device that may serve to focus expectations, as Schelling’s examples amply illustrate. Some asymmetrical solutions are recognized as such by two equally prudent players in a symmetrical dilemma, the disadvantaged party recognizing the prudence of asymmetrical play just as much as the advantaged party. Consider, for example, the following dilemma. You and two partners each receive one of the three letters A, B, and C. Each of you is to write the letters, A, B, and C, in any order. Only if the order is the same on all three lists do you get prizes totaling (say) one million dollars, half of which goes to the holder of the first-named letter on the three lists, one third of which goes to the holder of the second-named letter, and one sixth of which goes to the holder of the third-named letter. This dilemma is symmetrical, since the prospects are the same for each player. However it enjoys a natural solution for players who mutually recognize an alphabetical ordering: A-B-C. Why is this the natural solution — as even the holder of the letter C must, regretfully, agree it is? Because it enjoys a signalling power. The point is just that the prominence of a strategy combination is itself a legitimate, strategic, reason why a player might elect the relevant portion of it.

In a theory of strategy, the capacity to focus expectations is not merely a tie-breaking property of a solution. In some instances, a prominent strategy combination presents itself as the solution precisely because, and not merely in spite of, the fact that it does not qualify to stand among the most favored strategy combinations. For example, consider the following dilemma: two players are to choose, without communicating, between hearts, diamonds, spades and clubs; they receive prizes only if their selections match; if they agree on hearts, diamonds or spades, each receives $1000; if they agree on clubs, each receives $999. Clubs should be chosen — at least in typical instances of this dilemma — because it stands out, and because identical play will in typical instances be much more important to each player than the difference between $1000 and $999. It will be much more important to win either prize than to win neither.

The moral is that the theory of choice cannot appeal unconditionally to principles of symmetry or prominence. If it appeals to such principles, it must do so for the sake of coordination. And then the appeal (to symmetry, for example), must rest on the influence one expectation has on another. The appeal to symmetry rests on the
prophecy that one player's preference for another's performing a certain action gives that other player a reason to perform that action. Thus the appeal to some principle for the sake of coordination denies autonomy of the second kind.

IV. Prudence

What does prudence demand? My proposal is as follows. To be prudent is, among other things, to be susceptible to the clues that are capable of bringing expectations into convergence, when failure to concert action would be disastrous for all involved — even when such clues lie outside the premium structure of one's dilemma-instances. To meet this demand, an individual must vigilantly monitor potential clues for the direction of mutual expectations, particularly when the necessity for tacitly concerted action arises — that is, when communication is impossible, inconvenient, costly or unwise. This requires being — as we all are — entities that reveal expectations and become aware of others' expectations, particularly under circumstances when revealing and making note of expectations are in one's interests. And perhaps the best means of ensuring that one will be in a position to concert action with others in circumstances where failure to do so would be disastrous on all sides, is by being the sort of individual who reveals and makes note of certain types of expectations as a matter of habit. It is plausible therefore to suppose that prudence actually requires, and not merely permits, individuals to be subject (in their expectations) to the psychological influences exerted by the expectations of others, as these help to regulate one's own expectations. For coordination of expectations is possibly even more important to meeting individual aims than (for example) alertness to the potential for profiting from vulnerabilities or exploitabilities of others. After all, one can aspire to exploit only those less advantaged than oneself, but one can aspire to cooperate with everyone for mutual advantage, no matter how strategically placed one happens to be.

Thus prudence cannot be averse to strategies for deliberation that, under appropriate circumstances, permits the deliberator to be subject to influence by the beliefs, desires, and expectations of other deliberators. Under appropriate circumstances, therefore, prudent individuals may be constrained by beliefs and desires not their own, for the sake of coordinating action, when failure to coordinate would be disastrous for all involved. Deliberative strategies which meet this condition would be found desirable by players in advance of being faced with decision, so long as the world offers advantages to those who cooperate.

It is, as a matter of fact, quite customary for social beings like ourselves to be regulated by the social institutions which exist specifically for the purpose of standardizing expectations. And humans do as a matter of fact depend on their estimations of how other individuals are liable to deliberate and form expectations, and act as if their own expectations have an influence on the expectations of others. Thus expectations, as a matter of human sociology, are mutually dependent. For formation of mutual expectations does not occur in an environment in which each expectant self is unaware of the anticipations of others, or in an environment in which the anticipated expectations of one individual can have no impact on those of the individual anticipating those expectations. Rather, formation of mutual expectations occurs only in a community of individuals of whose intellectual lives certain aspects are known to be, and known to be known to be, open to public inspection — entities whose intellectual lives, as well as their visible lives, are intertwined. Aspirations
that others should be vividly aware of certain of one’s anticipations need not be altruistic; for public access to certain aspects of one’s intellectual life is in many cases in the service of self. It is these (sometimes implicitly, sometimes explicitly) revealed, and hence anticipated, expectations which furnish mutual grounding, each expectation for the other. In fact, one might even say, with august precedents to give one the courage to say it, that rationality is a species of anticipatability.

Sensitivity to influences exerted by desires and expectations of others will form part of any prudent decision-making; but individuals to whom such a sensitivity belongs will not be autonomous, according to the second definition, since they will be constrained, under certain circumstances, by beliefs and desires not their own. One advantage of the proposition that prudence calls individuals to be (under certain circumstances) subject to the beliefs and desires of other individuals, is that it has the potential to justify cooperation in such dilemmas as the famous Prisoners’ Dilemma, under conditions which permit the achievement of a collective perspective on decision (in other words, when the decision makers are able to sustain certain mutual expectations), even if the individuals in question are purely self-interested. Thus the proposition according to which rational deliberators need not be autonomous (in the second sense of the term) explains, at the same time as it justifies, the collective actions of individuals who share a perspective on their situation, and will thereby serve the social and political sciences much better than the Autonomous Deliberator assumption.

V. REMARKS ON JUSTICE:
ON THE USES OF DECISION THEORY

There is a long and venerated tradition in philosophy, running from Socrates to Rawls, of founding justice on prudence, if only for the sake of solving the problem of recommending the (favored) system of social institutions to all. This (contractarian) tradition has had its impact on international law, promoting the conception of a legitimate system of governing institutions as one which enjoys the consent of the governed, however self-interested they might be. Some philosophers of social science believe that this tradition in social and economic theory presupposes that prudence calls individuals to deliberate as autonomous entities. For example, Merrilee Salmon, in a review article, writes: “Many contemporary social scientists[...]

V. REMARKS ON JUSTICE:
ON THE USES OF DECISION THEORY

There is a long and venerated tradition in philosophy, running from Socrates to Rawls, of founding justice on prudence, if only for the sake of solving the problem of recommending the (favored) system of social institutions to all. This (contractarian) tradition has had its impact on international law, promoting the conception of a legitimate system of governing institutions as one which enjoys the consent of the governed, however self-interested they might be. Some philosophers of social science believe that this tradition in social and economic theory presupposes that prudence calls individuals to deliberate as autonomous entities. For example, Merrilee Salmon, in a review article, writes: “Many contemporary social scientists[...]”

Others write that social scientists assume only that rational agents act as if they were independent entities, giving preference to their own self-interested desires.

The enterprise of justifying systems of institutions to the prudent presupposes that prudence comes before justice, in the ontological order of things as well as in sociology. It presupposes, at any rate, that prudence is an individual-centered affair, as such requiring no recommendation to an individual. While it might be true, perhaps even by definition, that prudence requires no recommendation, it does not follow from this that prudence is an individual-centered affair. And obviously it does not follow that prudence demands autonomy.

What consequences do our conclusions have for a theory of justice? There are certain agreeable propositions to be drawn from the conclusion that prudence admits of a group perspective on decision. For example, it will not be so very hard in every circumstance to commend justice to the prudent. But we must remember that a group perspective is not always achievable:
it is achievable only if the group's solution to a dilemma is more advantageous to each than the sum of individual-centered solutions would be. But this might (who knows?) cover a large enough proportion of cases that we should not be overly concerned.

The point, however, is that the problem of justifying justice, or even cooperation, to the prudent is solvable only on a case-by-case basis. And this point may not be agreeable to those who labor in the field of justifying justice to all, conceiving of this labor as an a priori enterprise. For it bodes rather badly for the idea that a single system or even a single conception of justice can be advanced universally.

Those who conceive of political theory as an a priori discipline might, on being persuaded of the points made above, conclude:

State University of New York, Buffalo

Received October 23, 1996

NOTES

1. I shall use the two terms interchangeably, and in such a way that game theory — which specializes in problems of strategy faced by two or more "players," each of whom is rewarded according to what all players choose — is a species of decision theory.

2. David Gauthier, in Morals by Agreement (New York: Oxford University Press, 1986, 22), writes: "Practical rationality in the most general sense is identified with maximization. Problems of rational choice are thus of a well-known mathematical type; one seeks to maximize some quantity subject to some constraint."


10. "Moral Orientation and Moral Development," in *Women and Moral Theory*, E. F. Kittay and D. Meyers, eds. (New Jersey: Rowman and Littlefield, 1987), 19-33, 31. In the very Introduction to this volume, the editors describe the two alternatives depicted by Gilligan as follows: "A morality of rights and abstract reason begins with a moral agent who is separate from others, and who independently elects moral principles to obey. In contrast, a morality of responsibility and care begins with a self who is enmeshed in a network of relations to others, and whose moral deliberation aims to maintain these relations" (p 10). The implication is that these two conceptions of the moral life are incompatible.


12. Thomas Hill, for example, argues that the ideal of autonomy contains many parts which are compatible with fellow feeling for others, in "The Importance of Autonomy," *Women and Moral Theory*, op cit., pp. 129-138.

13. Philip Pettit, in "The Virtual Reality of *Homo Economicus,*" (Monist 78, 1995, 308-329, p. 309), calls self-centeredness what feminists have called, and what I am here referring to as autonomy of the first kind. However the term "self-centeredness" misnames, as Pettit himself recognizes (p. 311 ff.), the assumption that one's decisions are affected only by one's own beliefs and desires, since one's desires need not be self-regarding.


17. For example, the ranges of application for the principle of Dominance and the principle of Expected Utility are disputed. See M. Resnik, *Choices* (Minneapolis: University of Minnesota Press, 1987).


21. Schelling actually conducts informal experiments in which subjects face this and many other interesting dilemmas. He reports his results in *Strategy of Conflict, ibid.*, chap. 2.

22. We are to imagine that you and I — now prisoners, previously partners in crime — are taken separately for interrogation, and given the same two options. Each of us may either keep silent and thereby become susceptible to conviction of a moderate crime which carries a modest sentence. Or each may confess and accuse the other of a heavy crime. If each is accused by the other, each stands to receive a heavy sentence. But an unaccused accuser goes free, while the heaviest sentence of all goes to an accused but unaccusing partner. It will go easier with you if you confess, whatever I do. For if I confess, you are better off confessing too, to avoid the heaviest sentence; and if I do not, still you are better off confessing, since you would then go free. And for me it is the same. But if we two confess, it will go much harder with each of us than if we two keep silent. This dilemma structure is very well known, and enjoys a wide range of exemplifications (see Resnik, op cit.). I will not present the argument for cooperation in this decision problem, but will leave that to another occasion.
