Alan Biller

"Risk Theory and Partially Non-Consequential Decisionmaking"

.

I am honored to have been invited to speak. on the extension of Risk
Theory (RT) to decisionmaking involving non-consequential norms.
Non-consequentiality is quite consequential for the claims made for the
applicability of RT to decision problems that, as I will discuss below,
are not transparantly tactical or wholly instrumental. On the negative
side I will argue that as a first order theory RT is not normative or
even a complete heoristic in such contexts. Thus I take strong exception
to the common tendencies, A) to link risk theory with welfarism and
versions of utilitarianism, and B) to claim that with or without (but
especially with) such linkage RT is normative in both moral and rational
analysis, choice and action. Recognition of non-consequential norms
leads to a rejection of the point of view espoused by Arrow [Essays in
the Theory of Risk Bearing]: that the description of a consequence
includes all that the agent values," so that he will be indifferent
between two actions which yield the same consequence for each state of
the world." (P. 45) It also leads to a denial of identification of any
version of consequentialism (e.q. utilitarianism based on a weak Pareto
principle) with rational thinking. Thus, contra Harsanyi, ["Morality and
the Theory of Rational Behavior". in Sen & Williams, eds. Utilitarianism
and Beyond] it is false "that the emergence of modern decision theory has
made ethics into an organic part of the general theory of rational
behavior." (P. 42)

45

On the positive - admittedly brief and speculative side, - - I will suggest that while RT is not normative in many decision contexts, it can be employed indirectly in a meta-role, namely to elucidate the emergence of non-consequentialist norms. In short my remarks are questioning, tentative and informal. My stance is that of a user and observer of RT, not that of a contributor to the substantive theory.

Paradigmatic applications of RT to individual and collective choice have been in decision analyses of capital budgeting, economic allocation, and the study of contracts (where it links concepts of information, agency and risk transfer). No-one familiar with this work would deny RT's power and relevance. Nonetheless I suggest that RT is normative only in restricted domains, i.e., those in which the governing assumptions include utility maximizing objective functions (and in social choice contexts some form of a Pareto principle), and in which there is prior agreement about the irrelevance of rights, duties, obligations, as well as practical (decisive) means of ordering preferences, assigning probabilities, and characterizing the very decision(s) at hand. Put another way RT applies best where many of the most interesting questions don't arise.

I had best start by saying what I mean by RT, tactics and strategy. RT is initially a theory of individual choice under uncertainty in which it is assumed that choices are to be made among states of affairs which are their consequences. Though not required by the theory, it is also usually assumed that these consequences are objects of preference and that these preferences are ordered (usually fully ordered with respect to transitivity and equality or indifference). There are usually no prior

46

limitations imposed on the domain of preferences. In theory all consequences are comparable or fall in a single preference metric. Thus, as in Arrow, valuation of consequences by preference (hence utility) becomes an exhaustive valuational principle for all rational choice and action. For a single agent RT is identical with individual utilitarianism under conditions of uncertainity. This sort of consequentialism does not directly imply utilitarianism as a (the) theory of rational/moral social choice. Social utilitarianism requires two further principles, the first a technique of preference aggregation (usually but not necessarily unweighted summation) and the second a rule (usually a form of the Pareto principle) for ranking aggregated preferences. Social utilitarianism of unlimited scope raises deep problems about rights and personal autonomy. And even with restricted scope under certainity (but with specific preference ordering principles) it leads to very deep problems of collective agency. I an not here concerned with these problems since I think utilitarian preference aggregation procedures are strained even in straight welfare decisionmaking, and quite implausible except as limited heuristics in more complex social and bureaucratic contents. However it is worth noting that the jump from consequentialism to individual utilitarianism is made easier by the comfortable economic theory of revealed preference, according to which choices are among preferences and that what is chosen is what is preferred. Thus rational ordering over preferences and utilities defines rational choice. But the identification of choice with preference is hardly empirically supported except in restricted domains and in conflict-free situations. Revealed preference should best be

47

regarded as a hypothesis about choices. RT only applies as a first-order theory to special cases and as a second order tool it can at most rationalize the etiology of non-consequential decision criteria. Since it does not reconcile non-consequential criteria with consequential criteria, revealed preference cannot be salvaged by applying to second-order or "true" or "corrected" preferences. I reject revealed preference as both a generally accurate empirical observation and as an axiom of normative decision theory.

The distinction between strategy and tactics is that between relatively ill-and well-structured problems. If one ignores the agents involved there is no essential difference between them, e.g. a given decision cannot be neatly classified as being in one set or the other. But decisions are made in contexts by specific agents. What renders a decision strategic is whether for that agent the problem is well-structured vis-a-vis other decisions he could imagine making (or has made). When related decisions fall in a hierarchy of relative abstraction, one gets the classical examples of perspective: The general tactically deploys his companies, and artillery within a strategic plan. But to the platoon leader on the ground the actions of the company and artillery are part of his strategic uncertainity. Strategy and tactics always underdetermine each other.

Following an honorable tradition which stresses internal relations and rejects unanalyzable entities, I believe that every tactical decision can be recast as a strategic one (and vice versa). Nevertheless we so strongly desire to avoid unstructured problems that we habitually adopt a tactical outlook as the preferred point of view. The same desire also motivates us to supress non-instrumental or non-consequential criteria. we do not like to be made painfully conscious of our relative ignorance about how to define a decision problem, how to specify the decision alternatives, how to take account (let alone trade-off) multiple criteria, or about what is the causal nexus of the choice itself.

RT has direct normative applications in those genuinely tactical settings in which consequentialism can be assumed. A mundane example is pension fund management. Onc can take as givens the ability to define the scope of the problem and the relevant consequences, as well as the ability to trade off different preferences, and to describe consequences probablistically. I confess to having myself worked on developing stochastic control models here with the purpose of defining decision rules that are optimal given trade offs among preference weighted possible consequences.

Unfortunately it is all to easy to derail this normative methodology. For example demands to invest in socially responsible ways (or not to invest in certain countries), or to resist seductions of self-serving fund management, raise issues of morality and law that resist reduction to consequential form. It is not just that a consequentialist reduction is theoretically difficult or incomplete: the disputants reject the very premises of the reduction. For example consequentialists often cite both the elegance of a unified theory with a single ordering principle and plead for the ideal of universal rationality. I for one remain unmoved. Elegance is desirable but nct determining: I regret that all is not water as much as the next man. As for the unversality of rationality, perhaps it is only obtainable via a complex theory with multiple ordering principles. And even if not obtainable - if for example multiple principles are necessarily sometimes/somewhere locally inconsistent, universal rationality may still be a useful regulative ideal. I suggest that belief in simple order·na consequentialism owes more to oversimplified philosophical theories (e.q. of essential natures, projects of or hierarchically ordered functions) than to impirical observation. In such a situation, it is mere rhetorical posturing to defame one's opponents as "irrational" or "non-moral" or to take the fact of eventual choice as expressing preferences. When pragmatics force trade offs, the choice may not be morally preferred, or likely to be stable. The next time, moral resolve may be stiffer. But have preferences changed? What is harder - and itself a less well defined problem - is to understand the extra argumentative factors which supplement dialectic.

To return now to the second question, I believe that RT can help explain the emergence and stability of non-consequential standards. There have been several very suggestive recent developments. First evolutionary theory has been applied to cultural transmission by Cavalli-Sforza and Feldman who [Cultural Transmission and Evolution, 1982] demonstrate how there can be evolution of "traits that are learned by any process of nongenetic transmission, whether by imprinting, conditioning, observation, imitation, or as a result of direct teaching." (P. 7) The cultural selection process is independent of Dalwinian selection - they can operate in opposite directions. Once suitable cultural objects have been identified, then, it is in theory possible to study their evolution in a context in which they may have no significant effect on the Dalwinian survivability of the individuals or groups which select and utilize them.

Non-consequential norms or values are clearly culturally transmitted and modified over time. What is now understood as the formal universalizability feature of ethical claims is a clear mutation in a series of norms containing the lex talionis, the proto-statistical theory of the normal (i.e. natural), hence that of the functional and essential, the interpretation of the essential as the intelligible and rational, and (in Kantian and Hegelian philosophy) the rational made a condition of human autonomy.

For non-consequential norms to have survived so long they must be what, following W. D. Hamilton, ["The genetical theory of social behavior", I and II, J. Theoratical Biol 7, 1-16, 17-32, 1964] & Maynard Smith, [J. Maynard Smith, Gene theory and the evolution of fighting" in J. M. Smith, On Evolution, 1972] one can call evolutionary stable strategies. Informally an ESS is one which is resistant to the introduction of mutants. Ordinarily an ESS is a behavioral policy whose evolution can be modelled directly in RT terms based on the attempt to each individual to maximize its expected utility. There is no need for this behavioral "choice" to be conscious. Can, then, issue then is non-consequential norms, which we presume are evolutionary stable, be shown to arise through the sort of social interactions in which RT illuminates individual choices? Since norms typically have force where there are conflicts between individual and group satisfactions, i.e., where individuals can improve their payoff by acting in socially sub-optional ways can one use RT to explain how can a non-consequential moral ESS evolve in such situations?

A model formulation by Axelrod & Hamilton ["The Evolution of Cooperation", Science, V.211, 3/81, pp. 1390-96] shows that the answer can be "Yes". That is, a non-consequential norm can emerge, be robust against competing standards, and become environmentally stable in the context of an iterated prisoners dilemma. They demonstrate that if the survival of a strageqy is sensitive to the average value of the consequential payoffs won by individuals using it, then the simple policy of tit for tat appears to superior to all competitors so far tested. Tit- for tat is, I suggest, an abstration from the simple rule of lex talionis. Introduction of the policy seems to require no more than its formulation within a group exhibiting genetic kin-altruism, itself enplicable in straight forward inividual RT terms.

As they note "Once the genes for cooperation exist, selection will promote strategies that base cooperative behavior on ques in the environment. We are on the verge of an RT explanation of the emergence of the simplest ethical norms. All that is needed (I say quickly!) is the theory of their subsequent cultural evolution.

None of this justifies reducing the content of non-consequential norms to pragmatic utility maximizing behavioral rules in any simple way. All one can claim is: A) non-consequential norms have survival value to the extent they maximize social utility in the long-term: B) they need not alter preferences but may only change the probabilities of acting rationally on preferences that violate group utility (N.B. the groups needn't be very large). The nature of the evolutionary sensitivity of non-consequential norms to their consequences is an area worth serious study. Finally, if evolution occurs both physically and culturally, can a RT explanation of the latter extend to the essence of strategic thinking - the structing of ill define problems through "creative" use of analogy? In this context prior subjective probabilities of consequences would be defined by the inherited and transmitted "strength" of the various analogies employed, while the range of choices is limited by the range of mutability of antecedently available models themselves evolving cultural artifacts. I would like to think RT will eventually illuminate creativity, but think this is a long way off.