For The Sage Encyclopedia of Political Theory
Robert Adcock ed. (GWU)
Project begun: 8/8/08
Draft as of November 25, 2008: text (7340 words) + biblio (1400 words)
Criticisms and comments encouraged.

RATIONAL CHOICE THEORY

Joe A. Oppenheimer, professor
Department of Government and Politics
University of Maryland, College Park, MD 20742
joppenheimer@gvpt.umd.edu

Contents

RATIONAL CHOICE THEORY

Joe A. Oppenheimer, Professor

Here, in the early 21st century, rational choice theory has become *the* paradigmatic way of analyzing behavior. Three relatively independent fields have evolved with rational choice theory at their core: game theory, social choice theory, and decision theory. Rational choice theory's status has evolved as its traditional structure has been morphed by the analysis of data from experimental efforts to test its assumptions. These tests uncovered inherent problems in its assumptions and are leading to fundamental changes in the structure of the theory. Rational choice theory has deep roots in economics and has become the foil for the development of a cognitive theory of choice in psychology. It has made important inroads in political philosophy and in political science as a whole. In a nutshell, rational choice theory is three things at the same time: it is both a normative and an empirical theory of individual behavior, and also a formalized logical structure that serves as the foundation for much theorizing in political science and economics. Essentially, it ties individual choices to preferences, underscoring choice as teleological or purposeful behavior. It claims both that we *ought* to behave purposefully in accordance with our values, and that we *do* behave so, although not all rational choice theories encapsulate all three of these elements. As such, choice is explained by preferences forming an intuitively simple story with both explanatory and normative presumptions and implications. The essay begins with a short history of rational choice theory and moves to its applications in political theory.

Rational choice theory's role in political theory is built on dual foundational presumptions that explaining individual behavior is the key to understanding the functioning of political institutions and that these behaviors can be aggregated to understand the behavior of the group. These presumptions fit well with the ideological justification of democratic political systems and are usually referred to as "methodological individualism."

THE NATURE OF RATIONAL CHOICE

BRIEF HISTORY

Although the origins of rational choice theory may be a bit murky, its modern roots stem from the age of reason. Its pivotal intellectual position was secured in Thomas Hobbes' Leviathan (1651). Hobbes tried to explain the basic functioning of political institutions via individuals' choices. He conjectured choices stemmed from universally held 'appetites' and 'aversions.' The effort was continued by such illustrious figures as Francis Hutcheson, David Hume, Adam Smith, and later Utilitarians as Jeremy Bentham and John Stuart Mill. Others followed including many in economics. These works spawned what has come to be thought of as classical rational choice theory.

Adam Smith emphasized the potential social functionality of Hobbes' simplifying assumption of self-interest, famously asserting, in the Wealth of Nations (1776), "It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own

interest" (p. 119). Utilitarians went on to formalize the link between individual choice and social welfare via a reduction of moral content to an interpersonally comparable utility *numéraire* that was seen as also motivating the individual. Of course, at the time, there was also less of a divide between the normative and the positive claims in political economy.

The utilitarian program[1] was difficult to sustain and began to unravel when, by the end of the 19th century, preference was stripped of its presumptive interpersonal comparability. By disparaging the notions of inter-personal comparability of satisfaction, or utility, Pareto (followed by virtually everyone, at least in economics, e.g. Robbins, 1938) reduced the power of the utilitarian framework in moral matters. Indeed, most would come to say that all that was left of the utilitarian program was Pareto optimality: If one can make others better off, without hurting anyone, only then is there an indisputable possibility of improvement for the group. Disempowered in matters of distribution and redistribution, Pareto optimality is a criteria extraordinarily tightly related to both 'efficiency' and 'unanimity.' As such, Pareto optimality became an almost universally accepted normative, yet weak, standard.

Marshall, and other economists such as Samuelson, later reduced preferences to a generalized value structure which eventually was defined by its assumed logical properties that then served as the deductive engine of micro-economics and game theory. These formal properties have come to define classical preference theory. Hence rational choice theory (choosing in accordance with one's values or preferences) is that theory that was developed from what became classical preference theory.

Classical preference theory

Modern micro economics and public choice, as well as much of political science and political theory is based on what we might refer to as classical preference theory. In this conception, theorems are derived from the formal properties of preferences and in many theorists' eyes, the 'realism' of the structure is simply not relevant (Friedman, 1953; but see Nagel, 1961; Frohlich and Oppenheimer, 2006). In this perspective, preferences usually are asserted to have the following formal properties:

1. They are *pairwise* – peoples' preference judgements are made in pairwise comparisons.

2. *Completeness* – all alternatives from which one chooses are comparable. And individuals (I refer to generic individuals as *i* and *j*) are capable of, and do form, judgements as to whether one item (I refer to generic items as $\boldsymbol{x}$ and $\boldsymbol{y}$) is better than another (usually written $\boldsymbol{x}P_i\boldsymbol{y}$) or whether the two are equally good ($\boldsymbol{x}I_i\boldsymbol{y}$). When *i* judges two items to be equally good, one says that *i* is *indifferent* between them. Completeness implies that for any $\boldsymbol{x}$ and $\boldsymbol{y}$, either $\boldsymbol{x}P_i\boldsymbol{y}$, $\boldsymbol{y}P_i\boldsymbol{x}$ or $\boldsymbol{x}I_i\boldsymbol{y}$.

3. *Transitivity* – allows two pairwise relations to be inherited by a third pair in the following fashion: if the relation is transitive then if $\boldsymbol{x}$ relates to $\boldsymbol{y}$,[2] and $\boldsymbol{y}$ relates to $\boldsymbol{z}$, then $\boldsymbol{x}$ relates to $\boldsymbol{z}$. Hence both preference and indifference are presumed to be transitive: for example, $\boldsymbol{x}P_i\boldsymbol{y}$ and, $\boldsymbol{y}P_i\boldsymbol{z}$,

---

1. Which is often referred to as consequentialism.

2. Note that here the relationship must be understood to be 'directional.' That is, just because $\boldsymbol{x}$ has a particular relationp to $\boldsymbol{y}$ does not mean $\boldsymbol{y}$ has the same relation to $\boldsymbol{x}$. Think, for example of the relation 'mother of.' It is directional. If Jane is the mother of Bob, Bob isn't the mother of Jane. Nor is 'mother of' transitive.

imply $x P_i z$.

4. *Reflexive* – any alternative is as good as itself.

Together, these four properties imply that preferences are a relationship over things that individuals can 'order:' In other words, one checks all alternatives against one another, and forms a complete ranking, rather than just selecting the 'best' from the field. Each of these assumptions can be somewhat relaxed without creating enormous changes in the theoretical conclusions.[3]

These formal properties are then related to individual choice by adding three more assumptions.

5. *Maximization* – Individuals are presumed to always choose their most preferred alternative.

6. *Stability* – the preference orders are stable over time and scenarios.[4]

7. *Uniqueness* – individuals have but one preference ordering.

Together these seven properties imply that we can explain an individual's choice behavior by an understanding of their preferences and the alternative consequences of the choices from which the choice is being made. Nothing else is required.[5] Since preferences of each individual are well 'ordered' they can be represented by numbers (although, in this form, the numbers only have ordinal meaning). The 'axioms' allow preference theory to perform some of the same roles as 'utility' theory did for the 18th and 19th century Utilitarians. But with utility no longer presumed to be inter-personally comparable it is robbed of most of its moral power.

In the 1930's and 1940's Von Neumann (1944, chapter 2)[6] expanded these presumptions so as to cover choices of alternatives that were associated with probabilistic rewards. Calling such alternatives 'gambles' or 'lotteries' he presumed people had preferences over gambles. He then expanded the properties to show that they could reasonably imply that rational choice led to outcomes associated with the highest expected value. In other words, preferences regarding the outcomes plus the probabilities associated with the outcomes involved in any gamble were all that was needed to evaluate gambles. To do this Von Neumann presumed all the above assumptions plus:

V1. *Reducibility* – the form of the lottery makes no difference: only the probabilities of receiving each of the possible outcomes matters. One would be indifferent between gambles which 'reduce' to the same rewards and probabilities via the calculus of probability.

V2. *Continuity* – Consider a person who prefers x to y to z. Then a lottery can be constructed

---

3. These are the mathematical properties of an 'ordering' (see Sen, 1970). Sen discusses how these properties can be 'relaxed' in some detail.

4. Of course if there is no sugar some might choose a cola drink rather than a coffee. Preference independence requires that the scenario doesn't add or subtract elements that directly affect the value of the alternatives to the individual.

5. A twist on the theory was added by Samuelson (1938), when he suggested the '*revealed preference axiom*:' that observation of choices is sufficient to identify the preferences of individuals. Though not essential for our purposes, it is worth noting that this was one of the factors that led to the empirical examination of choice behavior that is forcing the abandonment of the classical view.

6. An introductory formal treatment of Von Neumann's utility theory is contained in Luce and Raiffa, 1957, Chapter 2.

between x and z that the person will find of equal value to the sure bet of getting y.

V3. *Monotonicity* – a person faced with two gambles that involve the same two alternatives, except that in one the person has a higher probability of getting the preferred outcome than with the other, that person always prefers the lottery with the higher chance of getting the more preferred outcome.

V4. *Substitutability* – A person is indifferent between any lottery and another which has elements which are of equal value to the individual were they but there, at the same probability.

Taken together with continuity, substitutability implies that all alternatives could be evaluated in terms of their value equivalent gambles made up of the best and worst alternatives being considered. Further, these then can be ranked in terms of the probability of receiving the best alternative.

Together, these four additional presumptions imply that individuals evaluate lotteries in terms of the expected value of the lottery. In other words, people have preferences over outcomes and are risk neutral. They are indifferent to the form of the risks by themselves, only caring about the calculable probability of receiving each particular outcome. For example, were there no decreasing marginal valuation of money,[7] a gamble of $100 with a probability of .01 and 0 with a chance of .99 will be found equally good to a dollar available without risk. These moves made a representation of preferences over all outcomes (sure-bets and gambles) possible both in geometric and numeric terms. Geometrically, continuity and decreasing valuation (see footnote 7) leads to indifference contours and the like. Numerically, it permits one to interpret utility numbers in more than an ordinal fashion. Indeed, the presumptions that permit one to analyze the value of lotteries in terms of their expected value allow one to map utility numbers on an 'interval scale.'[8]

Still, without a positive element, classical utility theory is a bit vacuous. No universal values are left, no interpersonal comparability: leaving the power of the analysis to rest on a simple teleological presumption of maximization in human behavior. Unfortunately for economics as well as for political science and political theory, more was needed. To gain power, self-interest was raised in many theoretical contexts to the status of an axiom. Further, the behavior of a group of individuals was to be understood as the aggregate of the self-interested behaviors of the individuals choosing within the group.

As suggested, the classical theory had considerable success, and considerable limitations. But as with other empirical claims, its longevity was bound to its empirical accuracy. And as the claims of theorists utilizing classical theory increased and required testing, virtually each of its assumed properties was subject to test. These tests were often inspired by the failures of extensions of rational choice theory to non-market behavior. Tests that demonstrated failures were followed up with detailed examination of 'what could explain' the erroneous predictions, and the trail of that

---

7. That is, every dollar is equally valued, regardless of how many dollars one has.

8. On any such (interval) scale differences between values are interpretable, but the zero point is fixed arbitrarily. Familiar examples of such scales are Fahrenheit or Celsius temperature scales and calendar dates. Equal differences between scale scores correspond to equal differences in the attribute being measured. So a difference between 30° and 35°, is of the same magnitude as between 45° and 50°, within either Fahrenheit or Celsius scales but the difference of 5° does not correspond to the temperature of 5°. The important thing is to notice that in utility terms, more than ordinal information is being given.

research has led to an entire field often referred to as 'behavioral economics.' Although 'behavioral economics' is a bit far afield for this essay, we need to understand a bit about the contributions that experimental tests made in demonstrating the limitations of the classical theory. This permits us to understand the directions taken in the new theories of rational choice.

Modern rational choice theory

As indicated, classical rational choice theory began to falter when tests of behavior in non-market situations were examined. Certainly, Samuelson's (1938) suggestion that we understand preference theory by its empirics let the cat out of the bag. But it was the early experimentalists that showed first the smoke, and then the fire. Kenneth May (1954), for example, showed through a simple survey of students that people frequently held intransitive preferences. But a whole field of experimental economics was born when Vernon Smith found it useful to use experimental type exercises in the teaching of economics after arriving as a young professor at Purdue in 1955. His subsequent experimental work explored the fit of the rational choice theoretic models with the realities he was able to bare in the laboratory. By 1979, in pyschology, Kahneman and Tversky showed that the problem of intransitivities was much wider than May suggested.[9] The foundation for an empirical understanding of rational choice was laid.

Others ran experiments to show that the classic notion of self-interest was suspect (Frohlich, et. al., 1984; Kahneman et al, 1986).[10] Getting past self-interest meant considering what forms of other-regarding behavior that could be developed to explain the observations (Fehr and Schmidt, 1999; Frohlich, Oppenheimer and Kurki, 2004; Bolton and Ockenfels, 2000; Charness and Rabin, 2002). New forms of preference theory cropped up that involved other aspects of preferences that had been shown to violate the classical view, allowing for inconsistent choice over time, framing effects, probabilistic preferences and more. The question became how spare could the theory be to derive the hypothesis needed for research (see Bendor, 1995).

But the most fundamental alternative to the classical model of rational choice was developed by a couple of polymaths working in evolutionary biology and developing game theoretic models: Maynard Smith (1973) and George Price. Worrying about biological outcomes understood as outcomes of a long repetition of a type of interaction, they conjectured that the interactions that were more successful would lead to more 'rewards' and a higher probability of off-spring survival, thereby generating an evolutionary model of strategic development in a population. For them choice became probabilistic, and adaptable. And the models that were developed, referred to as evolutionary game theory, changed in both form and foundation, from what we earlier called classical rational choice theory to what has become the theory of evolutionary games. The introduction of this to political science was carried forward by Axelrod (1984).[11] In his work, the interaction of choice that was seen as repeating and over which evolution developed in most of this

---

9. This led to a large field of study in cognitive psychology called Prospect Theory. Numerous reviews of this literature are available. For examples, see Rabin (1988) or Quattrone and Tversky (1988).

10. This also led to a larger field of study that included what became known as dictator experiments. These are nicely reviewed in Roth (1995).

11. The state of modern evolutionary game theory is easily found in text books such as Gintis (2000).

analysis was that of a two-person prisoner dilemma game.[12]

THE FRUITS OF RATIONAL CHOICE THEORY

Noting that rational choice began in political theory but developed through the work of very diverse contributors, it is not surprising that the beneficiaries of these theoretical formulations were widely dispersed in such different disciplines as philosophy, economics biology, psychology, and political science. What follows, is an examination of some of the highlights of classical rational choice contributions to normative and positive theories of politics.[13] They include improved understandings of

- social cooperation through what is usually referred to as the logic of collective action;
- the behavior of collective actors (for example unions and governments) through what is usually referred to as social choice theory, and spatial modeling
- what might constitute a metric of social or collective well being and hence a yardstick for political performance.

RATIONAL CHOICE & POLITICAL THEORY

The classical view of rational choice has led to numerous contributions to both positive and normative political theory. These contributions have led to both further testing and further problems for the theory. In what follows we look at a few examples of the seminal econtributions of rational choice theory to the understanding of politics.

COLLECTIVE ACTION AND THE POLITICAL CONTRIBUTION PROBLEM

If shared interests are to be satisfied, and if satisfying them for one member of the group satisfies the others, then why would rational, self-interested individuals work to get their shared interests satisfied? This is a classical puzzle[14] that Mancur Olson (1965), Russell Hardin (1970), and others set out to solve. What emerged was a somewhat complex, but sobering view of how humans solve what might be called the collective action problem. The first steps in the modern puzzle solving were taken by Olson, who identified the basic conflict between self-interest and any 'natural coming together' of individuals to solve group problems. Hardin then recast Olson's relatively

---

12. This has, subsequently been generalized (see Bendor and Swistak, 2001).

13. Although there have been some interesting contributions of the post classical school, most of the work in this area is still over the foundations, and has not yet settled down to an aggregate, consistent, theory of political interaction. But see some of the chapters in Gintis, 2000 for a different perspective.

14. Jean Jacques Rousseau was but one of numerous classical authors who analyzed this problem. See his Discourse on Inequality.

complex argument as a simple n-person prisoner dilemma game.[15] This moved the analysis along greatly in two ways. First, it gave us a number of rich implications to interpret what to expect in our political behaviors. Second, the implications permitted experimentalists to develop an easy set of models to test.[16]

Key Implications of the Logic of Collective Action

As already pointed out, in a prisoner dilemma the individual has a dominant strategy to not contribute. The individual, in an 'unorganized' group won't contribute on the basis of the public good alone. This led Olson to note that the incentives that lead one to contribute toward the socially desirable supply of a public good must be both somewhat independent of the good and work in an individualized fashion. As Olson pointed out, unorganized groups won't get their shared interests met except as the externality of other activities. But how and why do groups get organized? The shortfall of the unorganized, or non-cooperative, outcome that occurs in the group can be roughly thought of as the difference in the aggregate values of the cooperative and non-cooperative outcomes for the members of the group.[17] This gain is the maximum that the group could spend on organizing and still have a net benefit from the organizing effort. Or, this net gain is what the group's organizers can tap to improve the state of the group, and reward themselves for the effort (Frohlich, Oppenheimer and Young, 1971).

If nothing else, our analysis to this point gives a solid justification for liberal political orders. Of course, there is no 'ought' derived without a normative presumption. In this case, the normative presumption, which we argue is inherent in the justification of democracy, is that it is a good thing for people to get their shared needs met (perhaps understood as a minimalist form of consequentialism[18]). *If* we agree that people should be able to meet their shared needs (and most of us do believe this) then it follows that people ought to have such basic freedoms as press, speech, and assembly. Without such freedoms, even the identity, and certainly, the aggregate value of the shared interests will likely remain unknown.

To justify government by improving the lives of the citizens, given the lessons of the logic of collective action literature, implies that people must be given basic civil liberties. Otherwise, the demand for many valued public goods will neither be manifest nor factored into public decision making. This proactive justification for liberties goes beyond a more traditional justification, which

---

15. A prisoner dilemma game can be thought of as a situation where each individual would individually be better off (i.e. has a dominant strategy) not cooperating but everyone would be better off if they all did (and therefore even had to) cooperate. Other forms of games have been found useful in analyzing families of social and political interactions including chicken and assurance games (see Dixit and Sheath, 2004).

16. See Ledyard (1995) for a comprehensive, but dated, review of the experimental literature. He points out that the 'environmental' variables most likely to have sizeable effects on the outcome are communication and the relative cost of a contribution to the public good. In other words, size, gender and educational makeup of the group don't make a significant difference to the outcome.

17. For, at least in some rough cut fashion, the aggregate value of a public good to its potential consumers is the sum of the values each of those consumers places on the good.

18. Consequentialism is the notion that an arrangement's value or worth is to be judged by its consequences. The most well known form of consequentialism is Utilitarianism.

turns on the need for "negative" protections from governmental intrusion. Generally, individuals will not know that they share interests if they do not have the possibility of free communication. This has been made apparent by how the Internet's low communication costs have led to greater awareness of the shared interests of such groups as gays and lesbians and other previously oppressed individuals. For groups to demonstrate the scale of their demands socially and politically, they must be capable of sharing the costs of the political efforts to change the public policies underlying their demands for public goods without undue costs being imposed upon them because of their identity. Since a means of sharing the cost of a public good is organizationally difficult, for groups of people to meet their shared needs over time they must have the freedom to organize themselves politically.

The size of the group doesn't change the behavioral prediction in a n-person prisoner dilemma. But, if interests are shared, the value of their achievement increases. Hence, if collective action is to be roughly understood as a n-person prisoner dilemma, solving the shortfall of a larger group is more valued than that of a similarly situated, but smaller group. Or organizing a larger group is more 'profitable' than organizing a smaller group. And if political leaders are somewhat oriented toward the private rewards they can gain from such efforts then politics is potentially more profitable in large groups and political competition will be stiffer in larger groups.

Along the same lines, we can note that the outcomes of elections are public goods, and that at least the instrumental value of voting (i.e. voting to change the outcome of the election) is likely to be very small since the probability of any one vote making a difference is very small. This conclusion has led to two inferences: First, voters will tend to invest little in acquiring information about political outcomes and alternatives, they will tend to be rationally ignorant.[19] Second, citizens not given reasons to vote separable from the effect of their vote on the outcome of the elections, are likely not to vote. The issue has been somewhat overdrawn in the literature, as many theorists argued as if there were no other reasons to vote.[20] In any case, the fact that many voters do vote, referred to as the voting paradox, has led to both an interesting research frontier and an Achilles' heel in the rational choice theory program (Green and Shapiro, 1994). It also led to a massive investment in experimental research to discover the precise limitations of the theory.

Key Findings from Tests of the Theory of Collective Action

The findings of the empirical tests did not fully confirm the predictions of the prisoner dilemma games. Finding individuals contribute more than predicted to public good problems, researchers such as Elinor Ostrom, and Charles Plott began to explore aspects of institutional structures that contributed to success in the sustaining of common property assets. In a dramatic series of experiments, Plott (1983) showed that the incentives generated by institutional design determined a great deal about the obtaining of shared group outcomes. Working on common property problems that Hardin (1968) believed required privatization of publicly possessed assets, Ostrom similarly found that institutional and environmental details made all the difference in sustainability: When deviant behaviors were easily monitored, and when punishment for non-cooperative behavior was easy, groups solved their public good problems (Ostrom, 1990 and Ostrom, et. al. 2001). From these discoveries grew a vast prescriptive literature that has had an enormous impact on institutional

---

19. The phrase is from Anthony Downs (1957).

20. But see the work of Frohlich, Oppenheimer and Young, 1978, where we argue that one might care about the size of the margin of victory or defeat, or more recently the similar, but broader, line of analysis by Gerry Mackie, 2003.

and policy design. It has led policies to use such market institutions as trading and auctions in such diverse problem areas as environmental, transportation, communication and other policy matters.

But the contributions of rational choice theory to institutional design[21] did not begin, nor end, with the study of collective action. It was a field that began with the negative findings of Kenneth Arrow who wondered which normatively attractive properties could one guarantee by properly crafting a constitution. The initial answer was astounding.

COLLECTIVE ACTORS, SOCIAL CHOICE, SOCIAL WELFARE AND THE 'ARROW PROBLEM'

Arrow's finding (1951) was fundamental to much of modern constitutional theory: if groups and group behavior are the aggregation of the choices of individuals and if individuals behave in accordance with rational choice theory, then democratic constitutions can't be designed to generate rational group choices. At the same time, the work implied there was a huge hole in our understanding of social welfare as the aggregate of individual welfare when we are bereft of interpersonal comparability (see above, p. 3).[22] The problems stem from the difficulty of using a rule to combine reasonable individual choices (or welfare indicators) in a manner that can insure a reasonable group choice.

We can identify some of the implications of his discovery by pointing to the problem of group voting cycles. Cycles lead to problems: a rule that leads to a cycle is aggregating a specific pattern of individual choices into differing group decisions depending upon extraneous factors. Simple majority rule exhibits this property when certain patterns of preferences underlie the choices made by individual voters. The cyclic results from those situations undercuts any simple assertion that group choice reflects the 'will of the group' or that it is somehow 'better' for the group than those that were rejected. The arguments force one to consider rebuilding the theoretical underpinnings that relate popular choice to notions of social welfare.

To illustrate a voting cycle, we construct an example: some voters with preferences over a set of alternatives and a defined voting procedure that lets us identify the winner. In this case the rule will be simple majority rule structured so that the voters consider their options two at a time. The winner of each contest (the one that gets a majority) survives to 'go against' the next undefeated option until only one option is left: the winner of the last contest. To show the problem of instability, assume:

> a group of 3 voters (*i*, *j*, and *k*) considering four options (***w, x, y, z***) with preferences of the following sort: two of the three, *i* and *k*, prefer ***x*** to ***y***; two others, *k* and *j*, prefer ***z*** to ***x***; and finally two (*i* and *j*) prefer ***y*** to ***z***. Further, let all three voters find ***z*** preferable to ***w***.
> Finally, we presume that ***w*** is preferred to ***x*** by *j* and *k*.

It ought to be clear, in such a case:

1. The outcomes would 'cycle:' Start with ***y*** : ***x*** beats ***y***, but ***w*** beats ***x*** ; then ***z*** can beat ***w***, but ***y*** beats ***z*** and we are back with the initial defeated alternative.
2. There is no 'stable' outcome.

---

21. For a good overview see Mueller, 1996.

22. This presumption is in keeping with the utilitarian perspective as outlined above (see p. 2).

3. Without further information, 'winning' or 'losing' in such a situation carries no normative weight.

With these preferences each of the options can lose: there is a majority that prefers some specific other outcome to each of the options. And each of the outcomes can win, including ***w*** (even though the voters unanimously prefer ***z*** to ***w***) depending upon the order of the contest. So the winning motion could be anything. Under these circumstances, what is chosen would be determined by something beyond preference, perhaps the order of the vote, or the structure of the agenda: in a legislature this is usually a strategic choice controlled in part by a committee chairperson, or a party leader.

Arrow and Impossibility

What Arrow shows is that majority rule isn't the problem. Cycles can only be avoided by rules that permit other things we won't like. For example, a dictator can be expected to choose in a manner that does not show this sort of instability but a dictator has other negative normative qualities. More specifically, Arrow shows there are deep conflicts between designing a social choice rule to achieve some minimal normative qualities that also requires avoiding a dictator. So if we insist that (labels for the properties are in parentheses):

1. Any pattern of preferences that the voters have can be aggregated into a decision by the rule (**U-universalism**).
2. Any side can win if it gets enough support (**positive association**). And certainly, this means that if the group is unanimous in preferring ***x*** over ***y***, the group choices ***x*** (**P- Pareto**).
3. The choice between any two alternatives should only reflect the preferences of group members over those alternatives (**I-independence**).
4. The results of the contests should be 'transitive' (which forbids the voting cycle problem identified above) (**O-ordering**):

These conflict. The first requirement permits us to choose any set of preferences (such as, for example the ones we specified in the example above). And the last two then can conflict directly, since transitivity (or a weaker version) would mean, for example that if the group chooses ***x*** over ***y***, and ***y*** over ***z***, then something else can be implied; with transitivity it should choose ***x*** over ***z***. But as we saw, this can violate the choice that would come from the a consideration of only the preferences of group members' over the final pair of options (***x***, ***z***). Which of these properties one would want to sacrifice is unclear.

For example, relaxing the ordering (transitivity) requirement does not come cheaply. This is because there is a connection between any notion of good and transitivity. After all our notion of good and better and best are related to notions of orderings: best is 'above' good which is'above' less good, etc. We need not insist upon full transitivity of group choice, for example. But the conflict in the conditions remains, just not quite as tightly. Sen (1970) shows that virtually any conceivable notion of best, or better, implies some sort of 'ordering' principle, such as transitivity. So when we say that one outcome is better for society than another, and that yet a third is worst, we are implicitly requiring some degree of ranking. Sen considers precisely what such normative terms might minimally require, and no simple solution allows for even loose rankings and **independence** (property number 3). Of course, we can also give up a strict interpretation of 3, and gain more

wiggle room as to what we can get from a choice rule (Aleskerov, 1999). All this has underlined that we have to think quite carefully about the normative properties we hope to obtain from our constitutions and the precise designs of our political institutions.

Challenging the non-comparability of utility and social choice

Of course, not all combined preference patterns lead to cycles with any particular rule. And not all decision rules lead to them either.[23] But those democratic (i.e. preference aggregation) procedures that do not lead to cycles implicitly involve some presumptive form of interpersonal comparability of preferences and utility. Were preferences totally non-comparable, this would be an arbitrary element of the rule that could deprive it of its ability to endow the result with normative weight. But precisely how relevant these are to the empirical problems of democratic choice is disputable.[24] Consider, for example, a comparison of majority rule with another rule: the Borda count. The difference that matters here is the amount of information that the voter gives in her ballot or vote: majority rule asks for very little information from the voter, 'What is your first choice?' The link to social welfare with majority rule must come from an assumption that the 'social welfare value' of each person's first choice (in any contest) is the same. If we can't compare the welfare, then at least we see which side leads to fewer disappointments.

With a Borda ballot, which rules out the simple voting cycle, the voter is asked to rank all the candidates. A higher rank is worth more points. If there are, say, 4 candidates, the top rank is given 3 points, and each subsequent ranked alternative is given one less: a 3rd place vote gives the candidate only 1. The winner is determined by adding up the total points that are given to any candidate, and the one with the most points wins. Of course, with Borda the preferences that determine the outcome between any pair will be the broader preferences held over the entire set of alternatives. But perhaps Borda can be said to do a better job than majority rule: after all, the voters are giving much more information about how the outcomes affect them. But there is still a need to consider how the votes being aggregated, relate to aggregate welfare. Now one voter's 2nd place counts equally with that of another's. And so on. This merely requires different assumptions regarding what interpersonal comparisons must be made to treat the aggregate Borda vote count as a legitimate indicator of social welfare: it doesn't let us avoid the need for direct comparison.

Spatial Models and Institutional Analysis

Another approach to the problem is to relax the first assumption in Arrow (**U**) and assume that individual preferences may have certain commonalities so that the decision rule need not generate a result for any mix of preferences. Specifically, Black (1958) has shown that if all voters agree on an underlying 'ordering' of the possible outcomes (for example from left to right, or guns to butter) and each voter has an ideal point in that ordering and prefers an outcome closer to her ideal than one further away, then there is no possibility of a cycle, or 'Arrow problem.' With most forms of simple majority rule, the social choice will turn out to be in agreement with the voter who is in the median position in the set of voters. Sen (1966) generalizes Black's single peaked preference argument a bit and shows that a somewhat broader set of conditions supports a median voter equilibrium.

---

23. For example, the unanimity rule will lead to a Pareto optimal outcome. But it can't help us get from a distributionally unfair outcome to one that reflects fairness or distributive justice. (See Sen, 1970).

24. See, for example, the dispute between Mackie (2003) and Riker (1982).

But Plott (1967) showed that when the dimensionality of political conflict goes beyond one, the simple story of a median voter equilibrium breaks down. But the importance of the median is interestingly maintained: only when there is a median in all directions do we get an easy equilibrium. Without that, McKelvey (1976) derived a 'wildness theorem' that showed that when a decision rule leads to cycles, the outcome might well lead anywhere. This led to a bevy of interest: after all, democracies don't seem to generate radical instabilities, nor do they seem to lead to 'randomly selected' inefficient outcomes (see Tullock, 1981, who conjectured that there must be some logic to the attraction of the center of the distribution of ideal points, perhaps via the obviously central positioning of the intersections of all the median lines[25] through the distribution).

Surprisingly, although there may be many issues dividing voters, measuring citizens' preferences often shows that the preferences of citizens are able to be mapped onto only one or two dimensions. So, for example, when there are but two contesting parties for public office, voters are often motivated by which of the parties' composite positions they feel 'closest' to (see Mackie, 2003, pp. 189-190). Hence Black's findings have relevance even in some substantially complicated electoral situations. But this wasn't totally satisfactory since the dimensionality of voters' preferences is not necessarily reflected in the legislatures and cabinets where laws are written.

Three major building blocks have been developed to justify Tullock's conclusion. First, it was shown that many 'institutional' structures can bring 'equilibrium' to an otherwise chaotic state (Romer and Rosenthal, 1978; Shepsle, 1979; Shepsle and Weingast, 1981). For example, many of the checks and balances of the American political system 'privilege' the status quo so that it becomes less likely to be caught in a cycle (Hammond and Miller, 1987). But far less elaborate arrangements, as shown by Romer and Rosenthal, can also generate stability. Second, developments have expanded the notion of commonality of preferences that can preclude cycles. Coughlin and Hinich (1984) establish that if the relationship between individual preferences and choices are probabilistic, electoral political competitions can lead to centralistic outcomes that are stable and conform nicely to conceptions of utilitarianism (also see the discussion in Mueller, 2003). Frohlich and Oppenheimer (2007) show that a consensus on a conception of justice can obviate problems of majoritarian cycles.

But clearly, one of the great advances has been the development of tools to analyze the dynamics of democratic politics in elections and legislatures when cycles do exist. Here the discovery of the uncovered set and the ability to identify its location has been crucial (Miller, 2007 and Miller et al.. 1989). The uncovered set might be considered to be the set of points that can be quickly recaptured (via a short cycle). Its location has been established to be within 4 radii of the smallest circle that intersects all median lines (called the yolk, see McKelvey, 1986). This has led to a more generalized understanding of equilibria put forward by Tsebelis (2002). But other advances were made to deal with such specific problems as stability in multiple party cabinet formation and multi-dimensional party conflict situations (Schofield, 1996; other models are discussed in Mueller, 2003, chapter 13). The upshot of all these pieces is that Tullock's conjecture is approximately right: there is considerable stability, centrality, and predictability to most democratic procedures. And this has led to increasingly successful modeling of specific political institutions and situations.

Social Welfare and the Evaluation of Governmental Performance

---

25. Such lines would have a majority of ideal points (including those on the line) on *both* sides of the line.

Not surprisingly, such a big problem was posed by the Arrovian analysis that it led to attempts to reconsider the link between rational choice and aggregate welfare (Riker, 1982; Mackie, 2003). Indeed, much of the effort of behavioral economists has been to investigate the link between "utility" (traditionally understood as welfare or happiness) and preferences. The discovery is that the satisfaction of individual preferences does not generally lead to more individual happiness.[26] Rethinking the problem of utility and its relation to preferences has led some to focus on more concrete aspects of preference satisfaction, such as the satisfaction of basic needs (see Braybrooke, 1987; Sen, 1999; Brock, 2005; Doyle and Gough, 1991; Oppenheimer, et al, 2008; and Oppenheimer and Frohlich, 2007). Rationalist branches of democratic political theory, led by John Rawls (1971) embraced such a substitution (referring to the goods that satisfied such needs as 'primary goods') even earlier than did the behavioral economists and social choice theorists. One of the ways that first Harsanyi (1953), and then Rawls (1972), moved beyond the consideration of preference was by reinstituting the notion of impartial reasoning in the exploration of individual and social welfare. This deliberate decontextualizing of choice forced its ties to something more basic than preferences.

These shifts from preference to more 'basic' indicators of welfare changed both the presumptions of inter-personal 'comparability' and the basic properties that one might expect of the individual welfare measures being aggregated into conceptions of social welfare. Shifting to such needs as nutrition, for example, invites one to compare the welfare of those who are hungry to those who aren't. Interpersonal comparability (Sen, 2002) is reintroduced.[27] The original difficulty with coming to grips with social welfare identified by Arrow is based on the emaciated non-comparability derived from considering preference satisfaction the be-all of individual welfare. Reformulating the epistemology of welfare reopens possibilities for tackling the old problems the Utilitarians and others explored. But now, we do so with a fuller understanding of how to consider what ought to be aggregated and what can and can't be obtained.

The tools of social choice were then expanded to consider these sorts of moral reasoning (Arrow, 1977). And the tools of experimental economics permitted the conjectures as to what are the contours and contents of impartial reasoning to be tested (Frohlich and Oppenheimer, 1992). Although there has not been a wholesale rejection of Pareto optimality as the criteria for governmental performance, there has been a resurgence of an investigation of the link between rational individual choice social justice.

Joe A. Oppenheimer

See Also. Condorcet, Marquis De; Bentham, Jeremy; Common Good; Enlightment; Game Theory; Hobbes, Thomas; Mill, John Stuart; Participation; Prisoner's Dilemma; Public Choice Theory; Public Goods; Public Goods Theory; Rational Choice Theory; Rationality; Rawls, John; Rousseau, Jean-Jacques; Scottish Enlightment; Sen, Amartya; Smith. Adam; Theories of Justice Neo-classical Economics; Pareto-optimality; Public Goods; Rational Choice Theory.

---

26. Most recently a whole field of 'happiness economics' has been established. Central to the research have been experimental and theoretical explorations: see Frey and Stutter, 2001; Bruni and Porta, 2005.

27. Pushing on the need to recognize the idiosyncratic nature of individual needs, Sen (1993) rejects 'basic needs' as a solid basis for moral theorizing. In its stead he, and Martha Nussbaum, have introduced the notion of capabilities (see Crocker, 1992).

BIBLIOGRAPHY

Aleskerov, Fuad (1999) Arrovian Aggregation Models, Kluwer Academic Publishers, Dordercht, 1999

Arrow, Kenneth J. (1951). Social Choice and Individual Values. (2nd ed, 1963) Yale: New Haven.

Arrow, Kenneth J. (1977). "Extended Sympathy and the Possibility of Social Choice." American Economic Review, 67 (December): 219 - 225.

Axelrod, Robert. (1984). The Evolution of Cooperation. Basic Books. New York.

Bendor, Jonathan (1995). "A Model of Muddling Through." American Political Science Review, vol. 89 (#4, December): 819 - 840.

Bendor, Jonathan and Piotr Swistak, (2001) "The Evolution of Norms." American Journal of Sociology 106 (6, May): 1493-1545.

Black, Duncan (1958) The Theory Of Committees And Elections, Cambridge: Cambridge University Press.

Bolton, Gary E. and Axel Ockenfels (2000) 'ERC: A Theory of Equity, Reciprocity, and Competition', *American Economic Review* 90(1): 166-93.

Braybrooke, David (1987). Meeting Needs. Princeton, N.J.: Princeton University Press.

Brock, Gillian (2005) "Needs and Global Justice." in The Philsophy of Need, ed. Soran Reader. Philosophy, Supplement No. 57: Cambridge: Cambridge University Press, 51-72.

Bruni, Luigino and Pier Luigi Porta (2005). Economics and Happiness: Framing the Analysis. Oxford University Press,

Coughlin, Peter and Muelvin J. Hinich (1984) "Necessary and Sufficient Conditions for Single-Peakedness in Public Economic Models, J of Public Economics, 25: 323-41.

Crocker, David A. (1992) "Functioning and Capability: The Foundations of Sen's and Nussbaum's Development Ethic." Political Theory v. 20 No. 4 (November): 584-612

Dixit, Avinash & Susuan Skeath (2004). Games of Strategy, 2nd ed. New York: Norton.

Downs, Anthony (1957) An Economic Theory Of Democracy. New York: Harper and Row.

Doyle, Len and Ian Gough (1991) A Theory of Human Need. Basingstoke: MacMillan Education Ltd.

Charness, Gary and Matthew Rabin (2002) 'Understanding Social Preferences with Simple Tests', Quarterly Journal of Economics 117(3): 817-69.

Fehr, Ernst and Klaus M. Schmidt (1999) "A Theory of Fairness, Competition, and Cooperation," Quarterly Journal of Economics, 114 Aug.: 817-868.

Frey, Bruno S. and Alois Stutzer (2001). Happiness and Economics, Princeton: Princeton University Press

Friedman, M. "The Methodology of Positive Economics," in his Essays in Positive Econcomics, Chicago: Univ of Chicago Press, 1953: pp. 3-43.

Frohlich, Norman and J. Oppenheimer, w Pat Bond and Irvin Boschman. (1984) "Beyond Economic Man." Journal of Conflict Resolution v. 28, no. 1, March, 1984: 3-24.

Frohlich, Norman and Joe A. Oppenheimer (1992). CHOOSING JUSTICE: An Experimental Approach to Ethical Theory. Berkeley: California University Press.

Frohlich, Norman and Joe A. Oppenheimer (2006). "Skating on Thin Ice: Cracks in the Public Choice Foundation." Journal of Theoretical Politics, 18(3): 235-266.

Frohlich, Norman, Joe A. Oppenheimer, and Oran R.Young (1971) Political Leadership and the Supply of Collective Goods. Princeton University Press: Princeton, NJ.

Frohlich, Norman, Joe A. Oppenheimer and Oran R. Young (1978), "A Test of Downsian Voter Rationality," American Political Science Review, 72 (March, No. 1): 178-197.

Frohlich, Norman, Joe A. Oppenheimer and Anja Kurki (2004) 'Modeling Other-Regarding Preferences and an Experimental Test', Public Choice 119(1-2): 91-117.

Green, Donald P. and Ian Shapiro (1994) Pathologies of Rational Choice Theory: A Critique of Applications in Political Science. New Haven, Conn: Yale University Press.

Hobbes, Thomas Leviathan (1651)

Gintis, Herbert (2000). Game Theory Evolving: A Problem-centered Introduction to Modeling Strategic Behavior. Princeton: Princeton University Press.

Hammond, Thomas. H. and Gary J. Miller (1987) "The Core of the Constitution" American Political Science Review. Vol 81, 1156 - 1174.

Hardin, Garrett (1968). The Tragedy of the Commons. Science, v. 162: 1243-1248.

Hardin, Russell (1971) "Collective Action as an Agreeable N-Prisoners' Dilemma" Behavioral Science 16 (no. 5): 472 - 479.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler (1986). Fairness and the Assumptions of Economics. in a specially edited volume (by Robin M. Hogarth and Melvin W. Reder) of The Journal of Business, 59 (No. 4, Part 2: October): S285- S301.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: an Analysis of Decision Making Under Risk." Econometrica 47 (March): 263-291.

Ledyard, John O. (1995). "Public Goods: A Survey of Experimental Research," in The Handbook of Experimental Economics, John H. Kagel and Alvin E. Roth, eds. Princeton University Press: Princeton. pp. 111 - 194.

Luce, Duncan and Howard Raiffa, (1957) Games and Decisions. NY: Wiley. (reprinted: NY: Dover, 1985).

Mackie, Gerry (2003). Democracy Defended. Cambridge University Press: Cambridge, England.

May, Kenneth O. (1954) "Intransitivity, Utility, and the Aggregation of Preference Patterns." Econometrica, vol. 22 (No. 1, Jan): 1-13.

McKelvey, R. D. (1976) "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control." Journal of Economic Theory, v. 12: pp. 472 - 82.

McKelvey, Richard D. (1986) "Covering, Dominance, and Institution Free Properties of Social

Choice." American Journal of Political Science 30 (May): 283-314,

Mill, John Stuart. "On the Definition of Political Economy, and on the Method of Investigation Proper to It," London and Westminster Review, October 1836. Essays on Some Unsettled Questions of Political Economy, 2nd ed. London: Longmans, Green, Reader & Dyer, 1874, essay 5, paragraphs 38 and 48.

Miller, Nicholas R. (2007). "In Search of the Uncovered Set," Political Analysis, 15 (1): 21-45.

Miller, Nicholas R., Bernard Grofman and Scott Feld, (1989). The Geometry of majority rule. Journal of Theoretical Politics, v. 1 #4 (Oct.): 379-406.

Miller, Gary and Norman Schofield (2008), "The Transformation of the Republican and Democratic Party Coalitions in the U.S." Perspectives on Politics, 6: 3 (Sept.): 433-450.

Mueller, Dennis C. (1996). Constitutional Democracy. Oxford University Press, Oxford, UK

Mueller, Dennis C. (2003). Public Choice III. Cambridge Univ Press, Cambridge, UK.

Nagel, Ernest (1963) "Assumptions in Economic Theory," American Economic Review (May): 211-219.

Olson, Mancur (1965) The Logic of Collective Action. Harvard University Press: Cambridge.

Oppenheimer, Joe A., Norman Frohlich, Maria Dimitriu & Cyrus Aghamolla, Some Democracies Are More Equal than Others: Using Social Welfare as a Metric for Political Evaluation, For presentation at 2008 APSA meetings, Boston

Oppenheimer, Joe and Norman Frohlich (2007). "Demystifying Social Welfare: Foundations for Constitutional Design." Maryland Law Review, Vol. 67, No. 1: 85-122.

Ostrom, Elinor, Robert Costanza, Bobbi Low, and James Wilson eds. (2001) Institutions, Ecosystems, and Sustainability. Boca Raton, FL: Lewis Publishers.

Ostrom, Elinor (1990) Governing the Commons: The Evolution of Institutions for Collective Action. New York: Cambridge University Press.

Plott, Charles R. (1983) "Externalities and Corrective Policies in Experimental Markets." The Economic Journal, 93 (March), 106 - 127.

Plott, Charles (1967). "A Notion of Equilibrium and its Possibility Under Majority Rule." American Economic Review, 57 (Sept.): 787 - 806.

Quattrone, George A. and Amos Tversky, (1988) "Contrasting Rational and Psychological Analyses of Political Choice." AMERICAN POLITICAL SCIENCE REVIEW. (82, NO. 3 SEPT.) 719-736.

Rabin, Matthew (1998). "Psychology and Economics," Journal of Economic Literature, vol 36 (March): 11 - 46.

Rawls, John (1971) A Theory Of Justice, Cambridge: Harvard University Press.

Riker, William H. (1982) Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice. Prospect Heights, IL.: Waveland Press.

Romer, Thomas and Howard Rosenthal (1978) "Political Resource Allocation, Controlled Agendas and the Status Quo," PUBLIC CHOICE 33 (4): 27 - 43.

Robbins, Lionel (1938). "Interpersonal Comparison of Utility: A Comment." The Economic Journal, Vol. 48, No. 192 (Dec.): 635-641.

Roth, Alvin, E. (1995) "Bargaining Experiments," in Kagel, John H. and Alvin E. Roth, eds. (1995) The Handbook of Experimental Economics. Princeton, Princeton University Press: 253-342.

Samuelson, P. (1938). A Note on the Pure Theory of Consumers' Behaviour. Economica 5:61-71.

Schwartz, T. (1981) "The Universal Instability Theorem," Public Choice, 37, no. 3, 487 - 502.

Schofield, Norman (1996). "The Heart of a Polity." in N. Schofield, ed., Collective Decision Making: Social Choice and Political Economy, Boston: Kluwer.

Sen, Amartya K. (1966). "A Possibility Theorem on Majority Decision." Econometrica 34: 491 - 499.

Sen, Amartya K. (1970). Collective Choice and Social Welfare. North Holland: New York.

Sen, Amartya "Capability and Well-Being," (1993) in Amartya Sen and Martha Nussbaum, ed. The Quality of Life. Oxford: Oxford University Press: pp. 30-53.

Sen, Amartya (1999). Development as Freedom. New York: Random House.

Shepsle K. & Barry Weingast (1981), "Structure Induced Equilibrium and Legislative Choice," PUBLIC CHOICE, 37, no. 3, 503-520.

Shepsle K. (1979), "Institutional Arrangements and Equilibrium in Multidimensional Voting Models," AJPS, (February, No.1) 23: 27 - 59.

Smith, Adam. (1776) The Wealth of Nations. New York: Penguin Classics, 1986.

Smith, Maynard J. and G.R. Price. (1973). The logic of animal conflict. Nature 246:15-18.

Tsebelis, George (2002). Veto Players: How Political Institutions Work. Princeton University Press. Princeton, N.J.

Tullock, Gordon (1981) "Why so much stability?" Public Choice.Volume 37, Number 2 / January: 189-204.

Von Neumann, John, and Oskar Morgenstern, (1944) Theory of Games and Economic Behavior. Wiley, NY.