



Amadae, S. M. "Rational Choice Theory." Encyclopedia of Governance. 2006. SAGE Publications. 19 Mar. 2011.

(uncertain) or well-defined (risky) probabilities of what outcomes may result as a consequence of actions.

Elements and Structure of Rational Choice Theory: Expectations and Game Theory

RATIONAL CHOICE THEORY

Rational choice theory is a theory of human action that is committed to expectations over probabilistic outcomes and game theory. From its original articulation, this theory of rational decision making was put forward as a new approach to economics, warfare, and social science more generally.

Rational choice theory is often simplistically considered to be a theoretical generalization of either technical instrumental rationality, requiring that an agent adopt the means necessary to realize a chosen end, or of economic efficiency, demanding the effective use of scarce resources as exchangeable means to achieve ends. Instead, rational choice theory represents a unique approach to social science that locates human rationality in an agent's mutually consistent hierarchy of preferences over all conceivable global nonexchangeable end states. Additionally, rational agents are presumed to make decisions (a) in strategic environments in which one agent chooses actions in direct response to the actions that others are calculated to take and (b) in situations with either unknowable

In rational choice theory, agents are described by their unchanging sets of preferences over all conceivable global outcomes, such as whether candidate Smith, Davidson, or Nelson will win an election, whether dinner will consist of chicken, fish, or tofu, or whether a public policy is one of waging war, negotiating a settlement, or relying on the international community of nation states to provide leadership. Agents are said to be rational if their preferences are complete, that is, if they reflect a relationship of superiority, inferiority, or indifference among all pairs of choices; and logically ordered, that is, they do not exhibit any cyclic inconsistencies of the sort: Chicken is preferred to fish, fish is preferred to tofu, and tofu is preferred to chicken. In addition, for choices in which the probabilities of outcomes are either risky, or uncertain, rational agents exhibit consistencies among their choices much as one would expect from an astute gambler. All of these consistency relations among preferences over outcomes are stated in mathematical axioms; a rational agent is one whose choices reflect internal consistency demanded by the axioms of rational choice. Rational choice theory holds that all considerations

pertinent to choice (that may include attitudes toward risk, resentment, sympathy, envy, loyalty, love, and a sense of fairness) can be incorporated into agents' preference rankings over all possible end states. Social scientists only have indirect access to agents' desires through their revealed choices; therefore, researchers infer back from observed behavior to reconstruct the preference hierarchy that is thought to regulate a rational agent's decisions.

It is generally not appreciated, but important, that the consistency constraints defining rational choice theory are not equivalent to those specifying maximization of marginal utility under a budget constraint, although formal bridging conditions may be added to achieve congruence. It is also the case that even though many social scientists that use rational choice theory adopt one canonical axiomatized form specified as "expected utility theory," the research paradigm sustains alternative views as well. There are subtleties in probability theory that divide researchers: Are probabilities objective features of the world or are they best regarded as subjective features of individuals' psychology, and what sorts of consistency conditions apply to decision problems that incorporate both attitudes toward risk and unknowable probabilities? The intractability of decision making in uncertain circumstances has lead to the formulation of bounded rationality that grounds rational choice in manageable rule-of-thumb calculations in a series of one-time circumstances. As well, psychologists have observed several prominent and predictable empirical deviations from rational choice theory that has made it possible to identify patterns of what may be termed "folk psychology."

Game theory, which relies on some form of expectations theory, provides a mathematical framework for analyzing individuals' mutually interdependent interactions. These agents are defined by their preferences over outcomes and the set of possible actions available to each. As its name suggests, game theory represents a formal study of social institutions with set rules that relate agents' actions to outcomes. Such institutions may be thought of as resembling the parlor games of bridge, poker, and tic-tac-toe. Game theory assumes that agents are like-minded, rational

opponents that are aware of each other's preferences and strategies. A strategy is the exhaustive game plan each will implement, or the complete set of instructions another could implement on an agent's behalf, that best fits individual preferences in view of the specific structural contingencies of the game. These contingencies include the number of game plays, the sequential structure of the game, the possibility of forming coalitions with other players, and other players' preferences over outcomes.

For social scientists using game theory to model, explain, and predict collective outcomes, games are classified into three groups: purely cooperative games in which players prefer and jointly benefit from the same outcomes; purely competitive games in which one person's gain is another's loss; and mixed games, including the prisoner's dilemma, that involve varied motives of cooperation and competition. Game theory is a mathematical exercise insofar as theorists strive to solve for the collective result of various game forms, considering their structure and agents' preferences. Equilibrium solutions are of the most interest because they indicate, following the Nash equilibrium concept, that given the actions of all other agents, each agent is satisfied with his chosen strategy of play. Equilibrium solutions have the property of stability in that they are spontaneously generated as a function of agents' preferences. Solving games is complicated by the fact that a single game may have more than one equilibrium solution, leaving it far from clear what the collective outcome will be. Moreover, some games have no equilibrium solutions whatsoever.

One perplexing feature of game theory relates to the assumption of reflexivity on the part of agents: Agents must choose strategies in response to their beliefs of what strategies others will choose. This idea of reflexivity leads some researchers to associate methodological individualism with game theory. This is the assumption that the individual is the pivotal unit of analysis for understanding collective outcomes in politics and economics. However, as the use of game theory for understanding interactions in populations studied in evolutionary biology makes clear, the assumption of reflexivity and a view of the individual that could sustain a liberal understanding of politics

and economics are not essential. Still, having made this observation, it remains the case that many who adopt game theory in social science find it consistent with individualistic approaches that view the individual as the sole determinant of personal preferences, goals, and values. It cannot be ignored that part of rational choice theory's outstanding successes in the late twentieth century is inseparable from its extensive refashioning of our understanding of how and why markets and democracy function to respect individual choices.

Applications of Rational Choice Theory to Problems of Governance

Bargaining, democratic processes for reaching decisions, the bases of social contracts, various constitutional designs, systems of incentives and punishments, processes for achieving conflict resolution, collective actions and the provisions of public goods, the assignment of rights, and distributive justice have all been studied using game theoretic models. Rational choice theory has become paradigmatic of social science because it has successfully navigated between explanatory and descriptive analyses of political phenomena on the one hand, and has provided useful tools for leveraging social scientific knowledge to better design institutions on the other hand. Thus, supposedly nonnormative findings from rational choice research have been applied to formulating public policies and to designing institutions. Much of the research within the paradigm with direct relationship to political economy and governance has been achieved in one of three schools: positive political theory, public choice, and social choice.

The first result derived from rational choice theory with clear implications for social welfare and democratic theories was the "impossibility theorem" derived by Kenneth J. Arrow in 1951. Starting with the assumption of individuals' rational preferences, the theorem proves that, given minimum conditions that many believe consistent with democratic will formation, regardless of what procedure is used, it is impossible to generate a collectively rational preference ordering over global social states. Given that democracy traditionally drew its legitimacy from claiming to deliver government of,

by, and for the people, the impossibility theorem is a setback for hopes that collective-will formation based on individual preferences can accurately reflect people's preferences. Similarly, the impossibility theorem challenges policymakers' ability to fashion public policies consistent with the public interest, as there is at this time no scientifically credible means to derive a comprehensive statement of the public good or social welfare from individuals' desires. The impossibility theorem thus served to cast the cogency of the paternalistic social welfare state into doubt.

In the 1950s, researchers exploited rational choice and the impossibility theorem to investigate further how individual choice leads to collective outcomes. Duncan Black developed the insight that under special conditions in which individuals' preferences exhibit the characteristic of being "single-peaked," that is, specially arranged from lowest to highest on one universal hierarchy between two poles, then collective-will formation can said to be a valid expression of individuals' interests. The "median voter" theorem, first stated by Anthony Downs, performs an analysis of rationally self-interested voters, finding that in running campaigns designed to win elections, rational candidates will cater to the average voter, as this mathematically ensures receiving the highest number of votes. In turn, William Riker demonstrated a feature the median voter theorem missed: That candidates' motives to cater to the average voter are limited by the extent they must reach to accommodate voters' preferences to establish a minimum winning coalition.

The insight that collective outcomes are best analyzed as the result of individuals' acting on rational preferences was also applied to the question of what types of constitutions such agents would select. Early research suggested that rational individuals would only agree to a constitutional framework as a result of unanimous voting that neutralized any citizen's fear that others' encroachment on personal interests could result. As well, rational choice research scrutinized the standard rule that collective decisions cast in accordance with majority rule are legitimate. Instead, it was proposed by James M. Buchanan and Gordon Tullock that rational citizens would uphold a greater-than-majority threshold for many legislative decisions

as a function of a cost-benefit analysis balancing the time and energy costs of reaching any decision against the costs of living under a government whose policies violate personal preferences. Rational choice was also used to define the problems of the “free rider” and collective action by demonstrating that rational individuals cannot easily cooperate to achieve mutually beneficial outcomes. Research on the behavior of rational voters raised the question of why individuals vote in the first place, given that there is only a minute mathematical probability that any single vote will affect the final election outcome.

More recently, the rational choice approach has been used to reconsider central questions in traditional political philosophy. The Western social contract tradition, relying on individuals’ consensual agreement to abide by contracts, has been reexamined using the idea that some multiparty, repeated form of the paradigmatic prisoner’s dilemma game is characteristic of the state of nature from which government and social contracts must emerge. Political theorists employing game theory revisit Thomas Hobbes’s seventeenth-century social contract theory put forward in his work *Leviathan* to understand how social order emerges from the state of nature characterized by a total lack of security consistent with each having the natural right to all things, including each others’ person and property. Rational choice theorists ask how it is that individuals can form a sovereign state given their character is governed by rational self-interest. As the prisoner’s dilemma suggests, whereas each person can see the prospective gains from cooperation, he or she has the ever-present incentive to cheat, either as a defensive tactic to avoid being the sucker, or as an offensive strategy to gain the most for oneself. It is widely held by game theorists that the prisoner’s dilemma captures tensions between individual action and collective outcomes that typify government: Each person calculates better personal payoffs by cheating the system or one’s fellow citizen, with the final result that each person is worse off.

Rational choice theorists disregard Hobbes’s social contract theory on the basis that it seems to presuppose what is in fact impossible: Agents can contract their way out of a prisoner’s dilemma game by

promising compliance with an agreement without an external enforcement mechanism to ensure compliance. As game theorists realize, Hobbes proposed an absolute sovereign that would enforce the social contract by coercive means to ensure compliance. However, it remains unclear how agents can establish a sovereign by mutual contract: If contracting were possible in the first place, then why the need for the absolute sovereign?

Instead of a social contract theory of government, rational choice theorists propose a coordination theory of government. Given that social contracts require third-party enforcement through incentives or punitive measures, and that such a third party cannot be presupposed in a state of nature, rational choice theorists argue that government emerged as a coordination game. In this understanding of social order, parties mutually realize that they are forever caught in the bind of the prisoner’s dilemma, with each poised to cheat the other. The only solution to the prisoner’s dilemma that is consistent with rational choice theory, leaving aside nonanonymous interactions in indefinitely repeated games more typical of small communities, is for all the parties to agree to establish an enforcement body to ensure individuals’ compliance with contracts. Thus, the contract itself is not a solution to the prisoner’s dilemma supposed to structure the state of nature. Instead, individuals’ mutual acknowledgement of the intractable nature of the prisoner’s dilemma is resolved through a more encompassing coordination game in which all parties accept the need for contracts combined with the need for sanctions to ensure compliance. As coordination games are defined by all agents having aligned preferences that reflect their agreement over ends, no recourse to promises, duty, or commitment to principle is required to understand the establishment and maintenance of government. Crucially, then, the social contract is disregarded as means by which stable government is secured.

Along with the social contract, theories of state legitimacy based on consent are also in disfavor among rational choice theorists because legitimacy requires a normative foundation that positive political theory cannot countenance. Instead, rational choice theorists rely on the concept of “mutual best-reply”

from game theory to assess whether state institutions are stable. Rather than relying on consent to a set of rules or constitutional principles to indicate their legitimacy, rational choice theorists look to individuals' revealed choice of actions as a clear demonstration of their preferences. In this view, the mere fact that individuals choose their own actions and, therefore, participated in the creation of a given collective end point, indicates their compliance in bringing about that end point. Even though a collective outcome cannot be evaluated for its legitimacy, game theorists do ask whether each agent would choose a different course of action if all other agents' actions remained constant. A stable, self-reinforcing equilibrium is achieved when every agent selects the same action in view of what every other agent selected to do. This is an informal statement of the "Nash equilibrium," which has become prominent for playing a role in political theory that resembles the role that consensual theories of legitimacy played in traditional social contract theory. The idea of consent to a set of governing principles is replaced by the idea that each person is satisfied with his strategy and outcome given what all other agents decide to do.

Rational choice theorists have reformulated the concept of trust, basing it on straightforward coordination, supplemented by sanctions, rather than principled agreement that may at some point in the future deviate from agents' assessments of their personal best interests. In this view, because it is irrational for any individual to go against personal preferences, trust among individuals must always be consistent with preferences in order to be a meaningful social category of engagement. Thus, for example, in a marriage, according to rational choice theory, trust among partners cannot be of the form captured by the traditional oath "for richer or poorer, for better and for worse, in sickness and in health, until death do us part." A viable marriage must, at all times, be consistent with both members' preferences, or it will result in at least one individual defecting from the marriage contract. Trust, in this view, is not predicated on loyalty or commitment through unforeseeable circumstances, but is instead based on moment-by-moment agreement with rational and unchanging preferences.

Other views of trust that invoke an agent suffering in violation of personal expected utility maximization are regarded as naïve and unrealistic.

Commutative and distributive justice have received copious attention from rational choice theorists. John Rawls's 1972 book *A Theory of Justice* was the most path-breaking attempt to develop a robust theory of justice to be wholly consistent with the idea of rational preferences. Rawls's attempt was communicating how useful rational choice theory may be for understanding the implications of constitutional principles of government. Whereas Rawls's first principle of justice, the priority of liberty, is familiar and protects individuals from external interference, his second principle of justice, the difference principle, is novel and was taken directly from game theory. Adapting von Neumann's argument, Rawls suggested that rational individuals, when deciding how to organize society, would adopt a strategy of minimizing the outcome that can be expected in the worst possible scenario in a multiparty game. Rawls maintained that the rationally self-interested individuals, who were selecting constitutional principles unaware of which specific role they would play in the resulting society, would only permit egalitarian institutions that guaranteed some positive benefit for society's poorest members. Although the inherent rationality of the minimax strategy continues to be debated, it is significant that Rawls and others believed that the entire project of constitutional design could be analyzed as a complex decision-theoretic problem that considers agents' anticipated outcomes in complex social interactions.

Conclusion

As an axiomatic treatment of rational human decision making, and as a method for studying collective decision processes, rational choice theory is studied in philosophy and mathematics departments, as well as throughout the social sciences in political science, economics, sociology, and psychology. The science of rational choice is both research on the abstract conditions, or norms, governing human rationality, and it also encompasses research that explains and predicts outcomes assuming rational agency. There are two

views on whether the theory simply represents a descriptive means to model behavior without presupposing that agents actually reason in accordance with the theory, or whether instead it actually describes the decision rules manifested by rational agency. Researchers upholding the first view are content to use the axioms of rational choice to model actions and predict outcomes. The second view maintains that rational actors exhibit purposive action consistent with the behavioral norms of rational choice. The first view is modest by not suggesting anything about the internal thought processes of agents; the second view upholds rational choice theory as a theory that describes the normative foundations of rational decision making. Even though the first view is more restrained, and is sufficient for applying rational choice methods to understanding social and political phenomena, many researchers hold the view that rational choice theory is a powerful analytic tool precisely because it reflects the actual principles that must characterize purposive agency.

Rational choice theory has been central to methodological debates throughout the social sciences because of its adherence to a limited view of human rationality as consistency among preferences that categorically deems irrational modes of conduct not reducible to this description. As with any robust research tradition, intense controversies abound both internally and externally. Debates internal to the field have tended to focus on complex nuances of the formal theory, as well as the suitability of associating consistency of choice with choices characterized by narrow self-interest. Whereas the former is previously touched on, the latter attempt, for example, is to determine if altruistic behavior can be consistent with rational choice. As it currently stands, researchers agree that altruistic preferences can be readily encompassed within rational choice theory, but this leaves open the question of whether a satisfactory concept of altruism can be reduced to agents' preferences over outcomes. Even though it seems widely recognized that any agent whose behavior fully resembles that predicted by rational choice theory would be either a mythical construction at best or a nonfunctional social idiot at worst, it also seems to be widely accepted that

at the current time there is no compelling alternative that better captures what many people now refer to as purposive agency.

Rational choice theory is advanced as a positive as opposed to a normative theory of social science because it obeys what many consider to be a canonical rule of scientific investigation: Build testable theories with observable facts, mathematical relationships, and uncontroversial minimalist assumptions. From its birth in the eighteenth century, social science is believed by many to stand in contrast to philosophy, morals, metaphysics, and religion, specifically because it studies humans as they are and not as they may, in some ideal world, be. Social scientists attempt to describe human agency as it currently exists and do not strive to alter people's underlying motivational rationales. This steadfast commitment among many social scientists to advance value-free theories of human behavior and collective outcomes is consistent with the abiding understanding of many since the Enlightenment, that the individual is the final and sole judge of her own ends, conscience, and rationales for conduct. Many researchers believe rational choice theory to be consistent with this distrust of normative theories of human choice and social formations that purport to tell people how to live their lives or govern their society.

—S. M. Amadae

See also Bounded Rationality; Collective Action; Communicative Action; Communicative Rationality; Equilibrium Theory; Game Theory; Impossibility Theorem; Logic of Appropriateness; Optimal Decision Making; Pareto Optimality; Positive Political Theory; Principal-Agent Model; Prisoner's Dilemma; Public Choice Theory; Revealed Preference; Satisficing Behavior; Social Choice

Further Readings and References

- Amadae, S. M. (2003). *Rationalizing capitalist democracy: The Cold War origins of rational choice theory*. Chicago: University of Chicago Press.
- Axelrod, R., & Hamilton, W. D. (1984). *The evolution of cooperation*. New York: Basic Books.
- Binmore, K. (1994–1998). *Game theory and the social contract* (Vols. 1–2). Cambridge: MIT Press.

- Elster, J. (1983). *Sour grapes: Studies in the subversion of rationality*. New York: Cambridge University Press.
- Hardin, R. (1988). *Morality within the limits of reason*. Chicago: University of Chicago Press.
- Hargreaves Heap, S., Hollis, M., Lyons, B., Sugden, R., & Weale, A. (1992). *The theory of choice: A critical guide*. Oxford, UK: Basil Blackwell.
- Hargreaves Heap, S., & Varoufakis, Y. (2004). *Game theory: A critical text* (2nd ed.). New York: Routledge.
- Luce, D. R., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York: Dover Books.
- Hausman, D. M. (1995). Rational choice and social theory: A comment. *Journal of Philosophy*, 92(2), 96–102.
- Mansbridge, J. J. (Ed.). (1990). *Beyond self-interest*. Chicago: University of Chicago Press.
- Mirowski, P. (2002). *Machine dreams: How economics becomes a cyborg science*. Cambridge, UK: Cambridge University Press.
- Monroe, K. R. (Ed.). (1991). *The economic approach to politics: A critical reassessment of the theory of rational action*. New York: HarperCollins.
- Rawls, J. (1972). *A theory of justice*. Oxford, UK: Clarendon Press.
- Satz, D., & Ferejohn, J. (1994). Rational choice and social theory. *Journal of Philosophy*, 91(2), 71–87.
- Sen, A. (1982). Rational fools: A critique of the behavioral foundation of economic theory. *Choice, measurement and welfare*. Oxford, UK: Basil Blackwell.
- Sen, A. (1987). Rational behavior. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: A dictionary of economics* (Vol. 4, pp. 68–76). London: Macmillan.