Rational and moral action: a critical survey of rational choice theory

de Jonge, J.P.R.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER XIV

RATIONAL CHOICE THEORY:
AN INSTITUTIONAL APPROACH

1. Introduction

The problems facing rational choice theory are partly empirical, partly theoretical in origin. The anomalies in expected utility theory were detected in empirical experiments, but the problem of free riding is connected with the construction of economic man as a selfish human being. The construction of atomistic agents creates difficulties in finding solutions to coordination- and cooperation problems and the independence axiom is responsible for problems in every situation where a choice is based on a comparison of states of affairs. That (parts of) theories are not confirmed by empirical facts is not unusual, and one can discuss whether the empirical tests were appropriate. When it comes to theoretical problems it is often the case that theoretical
constructs or concepts contradict each other. For instance, maintaining the independence axiom by individuating choice-options through broadening the definition of outcomes is not compatible with consequentialism. Another example is that the 'escape from psychology collapses with the 'escape from perfect foresight.' I have discussed these paradoxes and inconsistencies extensively in the preceding chapters and shall restrict myself in this concluding chapter of the Parts I and II to some general remarks about it.

In section 2 I present a short outline of the transformation of economic science in the previous century, followed by some important aspects of personal, strategic and social choice. In section 4 I present the main elements of institutional individualism. As usually, I finish with a conclusion.

2. The great transformation

In the development of economic science we can observe three 'revolutions'. These were all involved, in one way or another, with the fundamental question: should economic theory be based on assumptions about human psychology or on assumptions about rational choice?

The first was the marginalist revolution in which the model of man was significantly modified for the first time. (see Morgan, 2006) Whereas Smith pictured economic man as a mixture of talents, instincts, motivations and preferences, Mill's creation of *homo economicus* was a character whose emotional range was narrowed to economic motivations and propensities. Mill believed that there is only one constant motivation, namely the desire for wealth, accompanied by a dislike for work. (Morgan, 2006, 2ff) But Jevons painted economic man as a calculating consumer, whose motivations and actions are defined in unobservable psychological terms. Jevons 's economic man is a move away from Mill's *homo economicus* who accumulated wealth in the form of goods or money, towards an individual gaining enjoyment or utility from consumption of these goods. "Whereas Mill's picture of *homo economicus* still seems to refer to observable behavior
Jevons calculating man is an introspective character whose subjectively registered feelings we cannot access." (Ib., 12) Morgan concludes that "[t]he marginal revolution has effectively put the problem of economic choice at the center of economic behavior". (Ib., 18) Economics became the *science of purposeful action*.

The second revolution is the one that Robbins effectuated and in which economic science was transformed from a theory about efficient allocation in the ‘economic domain’ into a science of practical reason. Marginal analysis was no longer bound by the traditional economic domain of production and consumption, but was extended to all aspects of human life in which scarcity plays a role. And since scarcity stretches to the consumption of time, all activities are characterized by it. In the 19th century, during the marginal revolution, the conception of man's desires or his needs is primary and dictates his valuations and choices. From the mid-20th century under the influence of Robbins it is the valuations and choices that are dominant. Utility theory remained the theoretical vehicle for the economic approach to practical reason. It was still grounded in psychological theory. Robbins, for instance, eschewed behaviorism because it conflicted with purposeful human behavior. We know from introspection, he argued, that people order their preferences in a consistent way.

During the third, positivist, revolution, utility theory was stripped of its psychological content. This revolution was completed with Samuelson's theory of revealed preference, which was meant as a definite escape from psychology. Since the revealed preference theory one might make reference to human psychology in explaining economic behavior, but doing so was unnecessary. It is assumed that desires can only be maximized or satisfied by 'rational choices.' By making choices dominant over desires it was admitted that the economic agent could have any type of motivation, provided he chooses rationally. "The "reasons for action" are not the initial feelings of the subject, but are rationalized (or reasoned backwards) by the economist from looking at the
consequences. Model man (..) is a normative model of normative behavior for real people to follow." (Morgan, 2006, 21) Ultimately the economic science of practical reason was transformed in an axiomatic decision theory.236 The utility function was reformulated into a choice indicator or a preference ordering function.237

In the second half of the previous century the assumption of perfect foresight was questioned and ultimately abolished. The abandonment of perfect foresight paved the way for the learning process to become one of the central issues of neoclassical economics. The learning process clashed with the abandonment of psychic processes. There is an incompatibility between the willingness to pursue a pure formal approach that eliminates the study of psychic processes and the need to place the study of learning processes at the center stage. (Giocoli, 2003, 397)

To get the strong formal results, the individual rationality had to be combined with other basic tenets of modern economics -perfect competition and general equilibrium theorizing. But there is an issue for which the formal consistency does not suffice and that is the problem that agents differ in their beliefs and expectations as to relevant economic events. How can these heterogeneous agents be modeled in terms of equilibrating behavior? Neoclassical theorists repeatedly ran into the necessity of assuming perfect identity among the agents featuring in their models. It is no surprise that the conclusion was drawn that modern mainstream theory does not possess a theory of the individual. (Davis, 2002b)

236 Utility became an index of preference satisfaction instead of being an object of preferences.
237 Only when a complete ordering is assumed can choices be inferred from preferences. But usually choices just reveal preferences, and not the other way around. This creates a problem for the belief-desire model, for desires, instead of being reasons for actions, function in this way more like animating devices, as the Rationality Principle does in the model of situational logic.
3. Personal and social choice: a different approach

In this section I want to touch on a number of unrelated aspects of rational choice and emphasize the difference with mainstream rational choice theory.

In theories of rational choice it is usually assumed that agents take decisions on a case-by-case basis. But it conforms more to reality to assume that people have plans. These plans do not have to be worked out in detail, and they are not an endless sequence of means and (intermediate) ends, but they structure further reasoning about means and the like and they do this in ways that are shaped by the requirements for consistency and coherence of the intermediate ends. Plans are typically partial: they specify certain actions or ends, but leave certain issues about means, preliminary steps, specifications and the like to later practical reasoning. Relevant intentions provide a background framework for further practical reasoning. Plans do not necessarily serve only selfish ends. Of course, there are selfish actors but there are also altruistic actors and most actors sometimes act selfishly and at other occasions altruistically. This is as much determined by their disposition as by the normative expectations of relevant others.

Let us now turn to the question whether rational choice theory should be based on assumptions about rationality or about human psychology. Since, in rational choice theory, agency is more important than the individual we need not know everything about the agent. We assume that agents are rational rule followers who tend to be risk averse. We also assume that agents act on a plurality of motives and we need to combine our knowledge about human psychology with our knowledge about their perception of situations to understand the motive they acted on. This is a kind of behavioral or institutional approach Behavioral economics is an attempt to introduce into economic science some of the theoretical and methodological approaches of psychology (how decisions are influenced by loss aversion, anticipated regret, by perceptions of fairness etc). (Bruni & Sugden, 2007) In non-expected theory
these psychological notions have found acceptance. We have to construct decision theories that do not depend on the independence axiom in the traditional way and to accept that agents have non-separable preferences across alternative events. In order to make room for principled and moral actions we have, moreover, to accept that preferences can be process-nonseparable. These theories use a modified utility function, which consists of a term representing basic utility and a term representing an added, experienced, utility.

An approach to decision making that takes plans as point of departure dismisses the traditional case-by-case approach. Taken too literally this would mean abandoning most of game theory. But we could make a distinction between games without and with learning effects. This distinction carries over in the distinction between games that are played once only and games that are played repeatedly. Games that are played once only are similar to a case-by-case approach. This distinction carries through into interpretations of equilibrium points. Nash proposed two possible interpretations of equilibrium points: 'mass action' and 'rationalistic' solution. The distinctive feature of the 'mass action' interpretation is

---

238 Frey (et. al.) has suggested that people have preferences about how allocative and redistributive decisions are taken. They mention the right to participate in political decision making as a crucial characteristic of any democratic institution. Citizens may gain 'procedural utility' (as they call it) from such participation rights over and above the outcome generated in the political process. If people feel that society offers equal opportunities they may be less concerned with inequalities because they see that social mobility is real. Procedural and outcome utility are not perfectly separable. If a procedure is fair, an unfavorable outcome may have to be attributed to oneself. If it is unfair, one can blame an authority for the bad outcome. This attribution effect predicts a complementary relationship between processes and outcomes: fair procedures are more valued when outcomes are good. The net relationship between procedural and outcome utility then depends on the relative strength of the substitutive and complementary effects. (Frey, et. al., 2004). (see also section 4, chapter XVI)
the requirement that games are played repeatedly, though not per se infinite. The players are not modeled as perfectly rational, but rather as learning human beings. The second interpretation holds for games played once only. In order for players to deduce what other players will do, they must have full knowledge of the structure of the game. This second interpretation is closest to neoclassical theory and has become the dominant interpretation of NE. (Giocoli, 2003, 310 ff)

The main solution concept of non-cooperative game theory (the NE) is the keystone of the whole theoretical edifice of contemporary neoclassical economics. Today NE embodies the discipline's most fundamental notion, namely, that of rational strategic behavior. But the concept still lacks a compelling interpretation, neither the epistemic (which was conducted with Bayesian tools) nor the evolutionary approach is fully convincing. The epistemic approach requires the imposition of unrealistic conditions on the agent's knowledge endowments. The evolutionary approach does not solve the problem of equilibrium selection. Schelling's solutions to these problems were never followed. The reason was that he did not use mathematical tools (Rizvi) or because he assumed heterogeneous agents (Innocenti). Both points favor, in my view, Schelling's approach and I recommend using it in combination with experimental research for the emergence of rules in repeated games.

The introduction of learning processes and the re-introduction of psychic processes facilitates the discussion about more complicated forms of preference orderings. The use of single, global rankings is not very fruitful. It neglects the presence of lower-order and higher-order preferences, the use of meta-rankings and the framing of decisions. Agents are engaged in various social practices where they meet different constraints and expectations. As a consequence, agents 'frame' their decisions in different ways. This results in multiple, partial, rankings. When these partial rankings conflict too much, they may not fit into one global ordering. In such situations a meta-ranking, a ranking of these partial
rankings, may be applied. A meta-ranking will also be applied when an agent is engaged in strategic action and has to take into account what other agents are likely to do.

When we accept that preferences are framed, that next to outcomes procedures matter, and that consequences refer both to material and nonmaterial costs and benefits, then it is a straight case that preference fulfilment does not merely refer to mental states only. The ranking of preferences can as well refer to the expected satisfaction of a desire fulfilled as to the conviction that a certain goal should be realized (see chapter XVI). The ranking of preferences that incorporate incompatible values is based on ideals and intuition.

The motivating issue that inspired social choice theorists is: how can it be possible to arrive at cogent aggregative judgments about society given the diversity of preferences, concerns, and predicaments of the different individuals within society? (Sen, 1999). It was long taken for granted that the well-being of the individuals in society depended on the extent of the preference fulfillment for all individuals concerned alike.

When well-being is equated with preference fulfillment, then important distinctions -as between needs and luxuries or between demands for consumer goods and political rights- are lost. Human agents are approached as bundles of wants. They have numerous preferences, which can all be exchanged for they are all instrumental to the agent’s well-being. Well-being generated by the prevalence of human rights could be exchanged for well-being derived from other sources. Citizens are treated as consumers, who are sovereign in the execution of their preferences whether they buy food, luxuries, medical care or educational services.

In the social welfare function all preferences count equally. Economists influenced by logical positivism had come to the conclusion that interpersonal comparisons of utility had no scientific basis. This reduced the informational base on which social choice theory could draw. There was only one criterion for social improvement left and that was the Pareto
criterion. Since then theorists have sought for other kinds of information on which interpersonal comparisons could be based. There may be a case for taking incomes, or commodity bundles, or resources more generally to be of interest in judging a person's advantage and not merely the mental states they help to generate. This can be supplemented by also taking into account the differences in converting these resources into the capability to live well.

The literature on basic needs and their fulfillment has also provided an empirical approach to understanding comparative deprivations. We see an absorption of the former materialist' approach within the definition of well-being.239 (see Part IV)

4. The dynamics of institutional individualism

In Part II I discussed the unintended consequences of intentional individual actions. In the context of an individualist, decentralized regime it is obvious to use invisible hand explanations. Invisible hand explanations must ensure that social order is possible even when individuals act solely on their own behalf. Invisible hand explanations, therefore, must refer to some spontaneous evolved coordination device. The dominant approaches are those in which either the price mechanism or institutions are held responsible for the spontaneous coordination of the independent decisions of a multitude of individuals.

The explanation of an efficient allocation of resources by the price mechanism by means of deriving aggregate excess demand functions from the individual demand functions appeared to be problematic. The institutional approach did a

239 Cooter and Rapoport suggest a return to "the material definition of economics [welfare], an empiricist conception of method, and an interpersonal conception of utility." [i.e., to add non-preference based information, jpr]. (Cooter and Rappoport, 1984, 528)
better job. It explains the allocation of resources between markets and hierarchies, and introduced rule-led decisions. Because rules are generalized ways of doing things it provides as it were ‘rational expectations’. Thus instead of anonymous markets the institutional approach provides us with markets, organizations and institutions. This conceptual framework is appropriate to make the micro-to-macro transitions.

I presented an outline of an institutional approach based on institutional individualism. This is based on two premises. First, rules become only institutionalized when they are collectively accepted. This means that the embeddedness thesis and an institutional approach go hand in hand. Only by using an institutional approach in which actions are situated or embedded, can one explain why people obey rules; commit themselves to principles, etc. Second, institutions appear as exogenous variables and the implication is that institutions are both the conditions for and the consequences of intentional actions. Human actions are recursive, i.e., agents reproduce by their activities the conditions that made these activities possible. Routine is a basic element of these daily practices. The recursive character of social life is made possible by the constitutive rules that characterize social practices and the regulative rules that are made possible by continuous daily activities. Each activity that is repeated often enough will be cast in patterns. Habituation leads to role formation; institutionalization emerges when agents mutually regulate their patterns of routine. Institutions can be seen as embodiments of collective routines.

The recursiveness of social life finds its expression in self-reproducing systems.240 To illustrate this I refer to the

---

240 Hayek embraces, according to Zwirn, also an idea of the recursiveness of social life. "Hayek’s conception of tacit knowledge and learning presupposes a recursive organization of social structures and individuals. Neither social structures nor individuals can be thought of in isolation: each presupposes the other." (Zwirn, 2007, 75)
schedule hereunder that describes the micro-macro relationship. This schedule is borrowed from Coleman who used it to illustrate Weber's attempt to analyze the relation between the Protestant ethic and the development of capitalism. Weber was especially concerned with the relations (2) and (1); he was less successful in explaining the structures

(protestant ethics) (4) (economic system)

(1) │                                  ↑
(3)

↓

(individual) -----------(2)------------> economic values

(behavior)

(figure 1)

of capitalist economies on the basis of the individual orientations. (relation 3). (Coleman, 1987) In the 'non-functionalist' theories it was not allowed to explain macro variables by other macro variables (relation 4). Rational choice theory is mainly concerned with relation (2).

I have adapted the schedule to show the approach that follows from institutional individualism. We see a circular movement that runs from individual attitudes (habits), to

(social systems) -----------(4)------------> social structures

(1) │                                  ↑
(3)

↓

(individual) -----------(2)------------> individual attitudes

(practices)

(figure 2)

individual behavior (social practices), social structures and social systems. Together they form a complex of causal and
material conditions for action. Social structure and human agency are inseparable. The economic system supervenes on the behavior of the individual agents but it has acquired properties of its own, i.e., properties that cannot directly be inferred from the individuals who make up society. These properties constitute social structures. Social structures are properties of social systems and simultaneously reproduce social systems.

5. Conclusion

In Part II I discussed the unintended consequences of intentional actions in the tradition of invisible hand explanations. I preferred an institutional approach and selected institutional individualism as the institutional framework for rational choice theory. At the end of Part II I

241 Where I have adopted Giddens' definition of social structures as a matrix of rules and resources, Fleetwood suggests that Hayek perceived of social structures as the embodiment of the collective wisdom of society. "Economics, for Hayek, becomes an inquiry into, and explanation of, the various resources, mechanisms and structures that are drawn upon, reproduced and/or transformed in establishing economic order." (Fleetwood, 1996, 741) The reason that rules can serve as a basis for action is because they embody the collective wisdom of the society. "Owing to the existence of a set of deep structures in the form of social rules of conduct, a high degree of compatibility of actions and consequences is ensured." (Fleetwood, 1996, 744) The rules Hayek has in mind appear as a complex of rules of the laws of property, tort and contract, although equally important for Hayek are a series of tacit rules such as integrity, honesty, keeping promises, etc. This does not conflict with Giddens' view on rules.

In Fleetwood's interpretation Hayek is almost pictured as a predecessor of critical realism. In fact, Hayek never wrote about social structures, but always referred to institutions.

242 These emergent properties bring about macro domains that create their own dynamics. This opens the possibility for macro theories that are relatively independent of micro theories. For examples of research into industrial relations within the context of the interaction between industrial and institutional structures, with the latter being sustained by and sustaining the former, see Streeck (1992). For examples of relations see Akerlof (2007). For a behavioral approach to macroeconomics see Akerlof (2002).
want to refer to the possibility, mentioned before, to
distinguish two institutional variants of rational choice
theory. In fact, I want to take a step even further afterwards
and, referring to chapter II, make up these variants with two
distinct methodological approaches.

In chapter II I have drawn attention to the curious fact that
there appeared to be two models of rational choice theory,
that were engaged in quite different discourses, that never
referred to one another, that implicated two quite different
methodologies, and were still, in the eyes of authoritative
authors like Caldwell and Hands, in fact one and the same.
I have compared the model of Desire and Belief (MDB) and
the model of Situational Logic (MSL) and concluded that the
MSL is particularly suited to explain what happens in
situations in which the behavior of individual agents carries
no weight. I referred to strong competitive markets and to
evolutionary analysis (like that of Alchian). I called this the
competitive model of rational choice theory. This model is
further developed in neo-institutional theory as the theory of
generalized exchange and managed competition. This variant
is not about an aspect of human behavior; rather it is about the
economic way of life in which all aspects are reduced to the
economic aspect. In other words, it does rather describe a
market society than a market economy.

The goal of the other variant is the comparative analysis of
governance structures, which includes public governance
structures and hybrid forms. This is an extension of the new
institutional theory. I have called it the comparative model of
rational choice theory and it is the basis for the institutional
approach to rational choice theory that I apply in this book.

Human practices are socially embedded and, therefore,
infused with normative expectations. These normative
expectations reflect beliefs about what we expect other agents
to do. Rational agents know the social rules, and know how
and when they are to be applied. A theory of rational action
has to explain why people obey social rules, why they
commit themselves to principles and feel guilt or shame
when they fail to live up with them. A theory is not like an airline or bus timetable, as Coase once remarked. We are not simply interested in the accuracy of predictions. A theory also serves as a base of thinking. It must help us understand what is going on.

When I proposed to disconnect rationality and self-interest and to relate rationality instead to degrees of reflectivity, I created the possibility to question an agent about her/his choices and actions. A well-balanced description of reasonable behavior should be something more than a mere reference to someone's personal preferences. We do not want to know what he or she wants, but what he or she cares about, what his or her commitments are. Decisions have material and immaterial consequences that may interfere with status, principles, ideals, or convictions. Thus, the question to be considered is whether we can take into account, theoretically, how agents justify an action. Hodgson thinks that the clue is whether agents are sensible to “.. a motivation which both justifies a moral principle prescribing the action, and ‘excites’ or causes the agent to perform the action appropriate to complying with the principle”. (B. Hodgson, 2001, 284). Essentially the answer is to be found in the manner in which cognitive elements interact with connative processes in the determination of rational behavior (see Ch. XVII). Important, moreover, is the extent to which the individual acquires reflexive knowledge, i.e., self-knowledge, and develops his capacity for rationally controlling his (ultimate) wants and criticizing and modifying his passions. In this sense the agent acquires the capacity to be ‘self-determining’ of his actions (see the chapters XVIII and XX).

To make a start I shall analyze in Part III the relation between moral and rational action in an attempt to discover the demands that have to be imposed on actions which not only qualify as rational but also as justified.