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

de Jonge, J.P.R.

Citation for published version (APA):
CHAPTER I

RATIONAL CHOICE

1. Introduction

In this chapter I discuss individual decision making under the condition that the consequences of acts are (to a certain extent) known. This gives me the opportunity to discuss some general characteristics of rational choice theory.

I start with an overall look in section 2. In subsequent sections I discuss some of these characteristics in more detail. In particular I will pay attention to the change in the concept of utility, a change that transformed the concept of rationality into a formal concept, something like consistency (section 3). I shall furthermore discuss the question whether rational behavior is to be equated with self-interested behavior (the sections 4 and 5).

The individual approach to rational choice that is outlined in this chapter has among others been criticized for the reason that most social scientists are interested in the regularities of
behavior of many actors. Therefore, we should not focus on individual psychologies, but rather on the situational constraints that are faced by a collectivity of actors. We should turn to the theory of situational logic instead of focusing on a theory of intentional action. I shall discuss this theory in chapter II and compare it with the theory of intentional choice that is the subject of this chapter.

2. Rational choice theory: an outline

In general, intentional acts are explained in terms of the “real reason” agents have in performing them, where the real reason is understood to mean “the practical reason” which caused the act. In order to explain human actions in terms of reasoning processes that caused them, one must presume that in general people are reasonable. Why? Unless one assumes that agents are reasonable there would be no limit on the interpretations one could give of the agent’s psychological states. Any act could equally well be connected with any belief and any desire. Reasonableness and intentionality are inextricably bound up with each other. (see, Fay, 1996) “Irrational” behavior, therefore, is not intentional; it is ‘driven by blind forces.’

In rational choice theory human action is described as the upshot of two selection procedures. First, from all possible (and relevant) action alternatives those alternatives are selected which are feasible. Second, from this set of alternatives the preferred one is chosen. The selection of the feasible set is based on the (financial, legal, social, physical and emotional) restrictions an agent faces. On the basis of his preference ordering the agent makes his choice among the alternatives. It is assumed that the preference ordering is consistent and complete, which means that all alternatives are

---

2 Where I use here the term ‘reasonable’ others use the term ‘rational’. I want to express that an action must be intelligible in order to judge it. In order to be intelligible, an action needs not to be rational in the sense that it is the best thing to do considering the alternatives.
ordered on the basis of a pair wise comparison. An actor is acting rational when there is no alternative from the set of feasible alternatives of which the outcome is preferred over the outcome of the chosen alternative.

| restric- | preferen- |
| relevant | tions feasible | selected |
| alternatives | ----⇒ alternatives | ----⇒ alternative |

figure 1

Characteristically, the restrictions do not influence the preferences. Moreover, stable preferences are assumed.\(^3\) Changes in behavior are therefore usually explained by changes in the restrictions only. Furthermore, it is assumed that actions are motivated by the preferences (in common language, the desires) of the agent only. The expectations of an actor, with regard to the relation between action and outcome, are based on his expectations (in common language, his beliefs). The expectations of an actor give expression to his knowledge of the situation (the restrictions, the opportunities) and, as mentioned, his expectations concerning the likely results of alternative actions. They are a product of knowledge, experience and information. The expectations guide the preferences of an actor, but only the preferences motivate an actor to act.\(^4\) As Hume once put it: ‘reason is the slave of passions’. Reason is to be seen as an instrument for achieving ends that are not themselves given by reasons. Rational choice theory employs an instrumental

\(^3\) Usually, there is no reference to real preferences, but to ‘fundamental’ preferences, as health, wealth, status, etc.

\(^4\) I will use the terms of preferences and desires, expectations and beliefs as if they are interchangeable. But, in fact, the terms preferences and expectations are second-order concepts derived from the first-order or ‘folk psychological’ concepts desires and beliefs.
The concept of ‘rational behavior’ is a mixture of three, more or less separate, characteristics. The one is that it denotes ‘instrumental rationality’, the capacity to choose the right means (resources) in order to realize a given goal. The second element is that a rational agent has the capacity to allocate his scarce resources in such a way that he succeeds in maximizing his utility. The third element is that the agent is self-regarding. His motivation will always be that his well-being be served. I shall return to this issue; first I proceed with the preconditions of a maximizing choice.

In order to qualify as an optimal choice a rational choice has to satisfy some requirements. Two requirements are primordial: the amount and content of information to select the appropriate course of action should be optimal, and the ordering of preferences should be consistent and complete. I will discuss these requirements in turn.

The (subjective) costs of exploring options and figuring out what to do, form part of the burden of making a rational choice. Decision costs encompass the costs of acquiring, checking and processing information, of deliberating and of monitoring the action. What is the optimal amount of information an actor needs in order to take a rational decision? When the benefits are small, decisions costs may outweigh them. A rational actor, therefore, will continue collecting and processing information until the marginal costs equalize the additional benefits. However the benefits of
additional information can only be assessed after the information is acquired. Thus, this approach fails. An alternative view is adopted, namely, that a rational actor will make use of all available information. This is no guarantee for optimal decisions, though. Therefore it is assumed that rational actors will learn to collect information and will learn to process it until they are so acquainted with the professional perceptions that they will be able to avoid systematic errors.  

The discussion about the optimal information anticipates the much fiercer discussion about maximizing choice. It has often been the case in the economic literature that the rational agent is pictured as an agent with a comprehensive mind who knows exactly how to equate (marginal) costs and (marginal) benefits in every situation where he has to balance alternative options. For descriptive purposes the assumption of a maximizing choice merely indicates that a representative agent should, when confronted with two options, choose the better one. Maximizing behavior thus, expresses the idea that a choice is rational when the outcome of the chosen alternative is better than that of the alternative options that

---

5 This is a rather strong assumption for it is akin to the hypothesis of rational expectations. All economists certainly do not share this assumption, neither is it supported by ‘folk economy’, the common interpretation of economic events. Simon opted for a theory of rational choice in which the ‘global rationality’ of the economic model was replaced by a more limited, ‘bounded rationality’ in line with the amount of information that rational agents actually can process. He recommended that the commitment to substantial rationality be replaced by one of procedural rationality, wherein the latter is to be understood as behavior that results from learning by doing.

6 Conlisk (1996) discusses the arguments pro and con the assumption of bounded rationality in detail. Usually imperfect information is attributed to satisfying behavior. But the idea of satisfying, or bounded rationality, is not seen as an alternative to the rationality assumption, because it is considered to be just another formulation of maximizing behavior when one takes the decision costs into account. The only exception is made for situations in which the skills of the agent are overstretched.
are taken into account.  

The second condition for a maximizing choice concerns the ordering of preferences. Imagine an individual who has to choose between different baskets of consumer goods. A rational ordering of his preferences for these baskets presupposes: first, that all goods are exchangeable (this is the continuity axiom, that rules out any lexicographical ordering); second, if there is one more unit of some good in basket X than in basket Y, then the utility derived from the goods in basket X is greater than the utility derived from the goods in basket Y (this is the monotonicity assumption); third, the preference-orderings are consistent (the transitivity axiom) and fourth, all baskets are (pair wise) compared (the completeness axiom). Assume there are three baskets: A, B and C. The transitivity axiom says that when the agent prefers A to B (A>B) and B>C, then he should prefer A>C. The completeness axiom says that when there are two baskets A and B, then either A>B, or B>A, or A~B. Assuming there is no uncertainty, these three axioms are sufficient to derive a utility function (as a representation of a preference ordering). (see Hargreaves Heap, 1992; Hausman, 1992) When an agent’s preference ordering fulfills the requirements of a weak ordering, i.e., the ordering is a reflexive, transitive and complete binary relation, and when the ordering is continuous, then the ordering can be represented by a utility function, that makes his choice utility maximizing. The result does not indicate that an agent effectively maximizes. But this is generally taken for granted among economists.

---

7 Choosing the best alternative of those available is a question of optimization. "Maximization and optimization coincide when the ordering is complete, which it may or may not be." (Sen, 2000, 486)

8 This condition says that when there are two commodities in a basket, it is always possible to define a new basket, by enlarging the amount of one of the commodities and reducing the amount of the other that will enjoy the same preference as the original basket did.

9 Mongin argues that any account of optimization is prima facie open to an infinite regress objection. "A conveniently general formulation is to say that to optimize requires one to run a costly algorithm, that to optimally..."
The deliberative activities needed to rationally assess an action, are not required on each occasion. The idea that each activity is oriented towards discrete individual ends conflicts with practical reason. Activities are part of actions and actions are frequently part of broader plans. As a consequence agents will frequently reason from a prior intention to a further intention. "It is by way of plans (...) that the connection between deliberation and action is systematically extended over time." (Bratman, 1999, 30) How should we picture the kind of deliberation in which an agent is involved within a planning framework? On this occasion it is just rational for an agent not to reconsider some prior intentions and just to carry out his prior formed plan. By accepting a plan, an agent has committed himself (see Bratman, 1999). An action is the flux of self-monitored activity and is not, as is often supposed, a series of discrete acts, each of which is chosen self-conscious and aims at satisfying some want. The actor only grasps a purposive act reflectively. Thus, “the purposive content of every day action consists in the continual successful “monitoring” by the actor of his own activity.” (Giddens, 1993(1976), 89)

A last theme that has to be addressed in this outline of rational choice theory concerns the character of rational choice theory: is it a positive (explanatory) or a normative theory? Some theorists insist that rational choice theory is select the latter requires one to run another costly algorithm, etc." (Mongin, 2000, 95/6)

10 This conflicts with an axiom of expected utility theory, the separability axiom that says that each action should be judged on its own merits. Vanberg thinks that ‘case-for-case’ maximization is a paradigmatic condition for maximizing behavior in rational choice theory. (see Vanberg, 1994) But under the pre-condition of some modifications relating to motivation and intention, rational choice theory can accommodate plans. (see chapter XVII, section 5)

11 For simplicity’s sake I will often refer to actions as if I discuss discrete acts. I hope the reader will keep in mind that in reality actions are composed by sequences of single acts and plans may be composed of actions.
purely descriptive or positive. Its purpose is to predict human behavior. On this view rational choice theory describes certain regularities in the choice pattern of agents. Theorists who take this line also say that the concept of rationality has no real content. Rationality equals consistency. The question is whether choices are consistent with each other viewed from the perspective of the axioms of rational choice theory. Other theorists see rational choice theory as having a genuinely normative content: it tells us how, as rational agents, we have to choose.

Though the distinction between a positive and a normative view of rational choice theory is rather clear, it has become a confused distinction because their is some tendency to connect both points of view. The confusion is already apparent in the following remarks of one of the founding fathers of rational choice theory, Morgenstern. His argument is that "[The] theory, as formulated by the von Neumann-Morgenstern axioms, is normative in the sense that the theory is 'absolutely convincing' which implies that men will act accordingly. If they deviate from the theory, an explanation of the theory and of their deviation will cause them to readjust their behavior." (Morgenstern, 1979, 180). In Morgenstern's opinion, rational choice theory is not normative in the sense that it can serve as a reference for evaluating the actual behavior of people, but normative in the sense that people who do not act according to this doctrine, even after it has been instructed to them, act irrational. Harsanyi adds to the confusion by stating: "..., rationality is a normative concept: it points to what we should do in order to obtain a given end or objective. But, this concept of rationality does have important positive (...) applications: it is used for explanation, for prediction, and even for mere description of human behavior." (Harsanyi, 1986, 83) And even Elster offers no firm position, when he explains that rational choice theory "is first and foremost a normative theory and only secondarily an explanatory approach." (Elster, 1990, 19) 12 The

12 Elster thinks, that it is an explanatory approach if we might assume that
question concerning the character of rational choice theory, therefore, will keep haunting us. Of course it could be solved when we could say that rational choice theory provides causal explanations of actions, but also on this question opinions differ (see chapter II).

In the next section I will describe how utility theory was transformed from a theory about practical reason into a decision theory. In this process, the rationality concept became a pure formal concept referring only to a consistent ordering of preferences.

3. Utility, preference and welfare

A rational choice must be based on a preference ordering which complies with certain formal axioms. I already referred to these axioms (as transitivity, completeness etc) when I discussed a maximizing choice. In this section I describe how the concept of utility was deprived of all substantial meaning, and hence how the rational ordering of preferences only refers to a consistent ordering. As a consequence the concept of a rational choice has become a pure formal concept.

The classical notion of value was 'value in use' referring to a useful object. And in the overwhelming majority of cases the thing was designed and made to have just those qualities, which make it useful for a purpose, and it is said to have value in use in virtue of that fact. In the 19th century the classical notion of value in use was replaced by the notion of utility. Utility still meant usefulness, but now usefulness in a general sense. According to Bentham, by utility is meant that property in any object, whereby it tends to produce benefits, advantages, and pleasure and to prevent the happening of pain, evil or mischief. The 'principle of utility' is the principle that actions are to be judged by their usefulness in this sense. In this definition utility refers to usefulness in an objective sense. The subjective element is the mental process of people are rational in the normatively appropriate way.
evaluating an object's capacity of doing so. Bentham did not identify ‘utility’ and ‘pleasure’ directly with each other. These words refer to related but distinct concepts; a subjective feeling on the one hand and a property of things, acts or states of affairs, on the other. Bentham’s account of rational action is in terms of mental states that result from actions, and not in terms of desires alone. This gave reason a more active role. Reason could overrule passions, although only for the sake of achieving a greater overall balance of “happiness”. (see also chapter XVII)

In the material welfare school, particularly associated with Pigou, goods were seen as having utility if they contributed to a person's physical well-being. Physical well-being is objective, like the condition of one's health, as opposed to the subjective experience of enjoying a good meal.

At the turn of the 19th century the definition of utility was not so clear-cut any longer. Utility was still understood in the everyday sense as usefulness, but there was a rival definition in which utility was seen as the capacity of something to satisfy the desires of individuals. Pareto coined this ‘utilité économique’ or ‘ophélimité (ophelimity). "The difference between utility and ophelimity is thus the difference between 'useful' and 'desired"'. (Cooter and Rappoport, 1984, 515) This difference had important implication for policy-making. The material welfare school was convinced that the utilities of different persons could be compared, because it assumed a hierarchy of needs. A distinction could be made between necessities, comforts and luxuries and a redistributive policy, which raises the consumption of necessities would enlarge national welfare. But ‘ophelimity’ is related to feelings and is therefore genuine subjective.

In the early 20th century, economists began to despise any reference to utility, volition, pleasure and so on. The rigorous formal representation of the human agent prevailed over the realism of the representation itself. This was due to the dominance of logical positivism that eschewed introspection as a valid source of scientific knowledge. It forced terms like motive, preference, and desire to be discarded from economic
theory. In order to lie the foundations for a positive science all terms had to be freed from psychological connotations. Economists had to shift their analysis from the subjective motives of choice to the objective acts of choice.

Scientific method demanded to leave out of account anything that is incapable of direct observation. At first sight this seems very plausible, but on second thought it is doubtful whether it is really justified. Robbins argued that, after all, the business of economics is to explain certain aspects of conduct. And it is questionable that this can be done in terms, which involve no psychological element. And we do in fact understand terms such as choice, indifference, preference, and the like in terms of inner experience. He concluded "if we are to do our job in economics, if we are to provide a sufficient explanation of matters which every definition of our subject-matter necessarily covers, we must include psychological elements." (Robbins, 1932, 89)

The axiomatic utility theory, formulated by Hicks and Allen in 1934 and strongly supported by Robbins, replaced the desire-based concept of utility by a preference-related concept. It departed from classical utility theory in the sense that any connection with psychological hedonism was denied. Because the satisfaction derived from the fulfilling of preferences was considered to be ultimately subjective, interpersonal comparisons of desires were considered arbitrary. We may compare 'utility' differences between persons across states, but not 'utility' levels. This has become known as the 'ordinalist' revolution. In the preference-related conception of utility in scarcity economics, preferences for bread or for opera are on equal footing, whereas bread and music occupied different positions in the hierarchy of needs in the need-related conception.

The plausibility of the approach of Hicks and Allen still rested on some psychological conception of ordinal utility and for this reason Samuelson argued that they had not gone far enough to free economics from psychology. To accomplish this goal he introduced revealed preference theory. This approach allows the derivation of indifference
curves solely from observations of purchases in the market. Thus, the theory provides the essential link between individual demand functions (consumer choices) and preferences. When Samuelson proposed the revealed preference approach, he motivated it by arguing that, if preferences are to be a valid theoretical concept in economics, then they must be derived from objective data as prices and quantities. He claimed that certain assumptions about consumer choices capture the whole empirical meaning of utility theory. But the approach failed, because researchers did not succeed in deriving such functions from the available data. "In spite of Samuelson’s high hopes the revealed preference approach proved empirically useless". (Lewin, 1996, 1316)

In the preference-related conception of utility, utility functions became ways of representing preference-rankings. The relation between preferences and choices, however, is complicated. The assumption that preferences reveal themselves in the choices of agents created a circularity of reasoning in the explanation of choices. And in the case of interdependent choices, the choice of an individual agent does not have to reveal his real preference at all when he acts on a social rule. (Sen, 1973a) In the preference-based approach utility did no longer refer to the experience of happiness or

---

13 In some contemporary work the term revealed preference is used to address the question whether preferences can be inferred from choices given information about beliefs. Hausman reasons that a choice depends on both the agent’s preferences as his beliefs (i.e., his information and his ability to use this information). Given the set of beliefs, any set of choices may be consistent with any set of preferences. It follows that any inference from choices to preferences depends upon the specific premises concerning the agent’s beliefs. The term revealed preference is unclear and, therefore, should be abandoned. (Hausman, 2000) But revealed preference theory only holds that preferences can be defined in terms of choice regardless of belief, i.e., there is correspondence between choice and preference.

14 With the approach of revealed preferences, choices are not longer deduced from preferences; it is preferences that are revealed through choices. Within such an approach, rationality is no longer defined by reference to utility maximization, but is defined rather as sheer consistency.
pleasure, but preferences still indicated valuable things. This reference to value was lost in the modern formulation of the utility function by von Neumann and Morgenstern. Utility acquired the meaning: "the value of a function that represents a person's preferences." (Broome, 1991a, 3) Today utility theory is essentially concerned with establishing representation theorems. The notion of utility is deliberately not meant to refer to an experiential state. The general claim von Neumann and Morgenstern made with respect to utility is that it represents nothing more than a "quantity" or "number", with no "surplus meaning". They redefined utility as a preference-related entity, without incorporating a preference-satisfaction view. All that modern choice theory assumes is that people's preference-orderings conform to certain axioms: roughly, they are consistent.

But how can utility acquire meaning independent from (the satisfaction of) preferences? What remains, in this approach, of the relation between individual choice and individual well-being? If choices conform to a utility function that is defined in terms of a mere preference ordering, is there any relation between (individual) choice and (individual) welfare? Welfare economics assumes that conceptions of well-being are conveyed by preference satisfaction. This is a doubtful claim however. The choices made by an individual might or might not be directed to her good. If someone prefers smoking to non-smoking is her smoking then after all better for her; an indication of well-being, although it harms her health?

In spite of the abandonment of all references to subjective states, some authors keep on believing that, the greater the number, and the strength of a person's preferences that are satisfied, the greater that person's utility. Harsanyi says that the von Neumann-Morgenstern utility function indicates "...how much utility, i.e., how much subjective importance, he

---

15 Broome, for instance, emphatically denies that ‘utility’ and ‘good’ refer to the same thing. “(...) a person’s utility, as officially defined, has no necessary connection with her good.” (Broome, 1991a, 4)
assigns to various goals." (Harsanyi, 1975, 600) And the greater a person’s utility, the greater his personal welfare. The preference-satisfaction view is deeply subjectivist. To relate choice and welfare, it has to be. But obviously these writers put more weight on the individual’s assent to her own fate (to prefer smoking) than on her objective welfare (the risk of illness).

4. Rationality and self-interest

The question I want to discuss now is how far the construction of the human agent casts human beings as egocentric in their desires or preferences and what the consequences are for the rationality assumption. Usually the preferences of an agent are deliberately left open, signaling that they can be of all kind, though they must be self-regarding. Nevertheless, economic theory does seem to operate with a much narrower definition of preferences, they are not only self-centered or self-regarding, but also even exclusively self-interested.16

Economic science offers invisible hand explanations in which collective goods are attained on the basis of each following his/her self-interest. At the same time, economics is specialized in prisoner’s dilemma’s, which show that individual rationality collides with collective rationality. In prisoner’s dilemmas there is a clear precept of rationality: never choose a dominated strategy. And since there is, within standard economic theory, no place for collective choice, rational individuals will not act to achieve their common or group interest. Hausmann even puts it bluntly that “(..),economics studies the consequences of rational greed.” (Hausmann, 1992, 95) And "agents seeking their own well-

16 There is a deep ambiguity regarding the status of the self-interest assumption. This assumption should be seen as part of the construction of an 'ideal type' not as a real description of human beings. But both economists and laymen mix up these different functions of the self-interest assumption.
being are what make economics run, and theories, which dethrone this motive, cease to be economic". (Hausman, ib.) Elster, suggests contrasting ‘rational man’ (with consistent preferences) with ‘economic man’ (with selfish preferences). (Elster, 1983, 10)

Rational behavior, therefore, is often associated with self-interested behavior. The suggestion is that agents in their deliberations about the selection of an action alternative do not give weight to the interests of other people. (see, Elster, 1984 (1979); Hausman, 1992; Hollis, 1996) But there are also readings in which the self-interest assumption is weakened. Pettit presents a rather paradoxical argument. He argues that agents are not implicitly or unconscious self-regarding, as Becker argues (Becker, 1976, 7), but virtually so. In everyday deliberations and decisions there are certain alarm bells that make that agents take thought to their own interests. "Under the model of virtual self-regard, most actions are performed without self-regarding considerations, but that is true only to the extent that most actions do suitably well in self-regard terms." (Pettit, 2000, 42) But on another occasion he holds that economists picture human agent as individuals with "predominantly self-regarding desires." (Pettit, 1995, 314)

Basically, a person’s self-regard is served by success in those of his pursuits and relationships that he does not enter to improve the well-being of (anonymous) others. Self-regard indicates that an agent bases his choices on his own preferences without implying that these preferences are necessarily selfish. Many rational choice theorists endorse what is

---

17 Harsanyi distinguishes between internal and external preferences, a distinction that is quite analogue to the distinction between self-regarding and other-regarding preferences. He thinks it is just natural to define a person’s interests in terms of his personal (internal) preferences, with complete exclusion of his external preferences. (Harsanyi, 1988)

18 The distinction between psychological and philosophical egoism comes very close to the distinction between pure self-interest and self-regard or non-tuism. Psychological egoism deems agents as selfish, philosophical egoism merely holds that the only desires that can move agents are their own. (see, for instance, Hollis, 1996, 6)
known as non-tuism. This holds that people’s preferences are not affected by their perception of other people’s preferences. Utility functions are independent. Non-tuism and selfregard go together rather easily.

Parfit took quite another point of departure. He did not start with opposing self-regard to other-regard, but with opposing shortsightedness to self-care. He demonstrated that one could split the set of presumed self-interested actions into two disjoint subsets: truly self-interested actions versus the pursuit of present aims. What is truly in our self-interest is to take account of our foreseeable future needs. He claims that behavior that is labeled as self-interested is often more properly seen as behavior directed at present aims. According to Parfit there are only two tenable theories about rational behavior, the present aim theory and the theory of morality, for only these theories give to "I" and "now" the same treatment. (Parfit,1984, 148) Thus, whereas present aim theory is both agent and time relative, moral theory is both agent and time neutral. The present aim theory holds that agents care only about their current interests. To be coherent, this theory must also proclaim that agents are only concerned with self-regarding goals. No more than the theory of self-interest can it be successful at the collective level. Moral theory, on the other hand, must be successful at the collective level. Therefore, it is also concerned with other-regarding choices.

To put everything in perspective I conclude that rational choice theory does not assume that agents are selfish, i.e., pursue their ends without bothering about the costs they impose on others. It does assume that agents are self-regarding. Second, I propose to make a distinction between prudential concerns (self-care on the long run) and moral concerns (other-regarding care). This distinction fits in with Sen’s elucidating distinction between actions based on (extended) preferences (‘sympathy’) and actions that stem from a sense of duty (‘commitment’). In the case of sympathy one feels concern for the well-being of those people one is acquainted with, as members of the family, or friends and
neighbors. Their well-being has a bearing on one's own welfare, we could speak about connected (or extended) utility functions. In the case of 'commitment' one acts on a moral principle, even when this conflicts with one's self-interest, in terms of one's welfare. Commitment can drive a wedge between behavior and welfare. (Sen, 1977a) Thus, in prudential concerns other-regard is limited to relatives, friends and the like. Prudence is the core of utilitarianism, the ethical doctrine that is embraced by rational choice theory.

5. “Das Adam Smith Problem”

The discussion about selfishness, self-regard etc. has historical roots, which dates form the writings of Adam Smith. In his Theory of Moral Sentiments Smith argued that by nature man is first and principally recommended to his own care and, therefore, is much more deeply interested in whatever

---

19 On the basis of Parfit's point of view sympathy, as an agent-relative concern, is part of the present aim theory. But sympathy, as a reciprocal attitude, can also be part of self-care. Parfit's distinction is not without problems then. The distinction between present aims and concerns with a longer time-horizon could more fruitfully be used to distinguish lower-order and higher-order preferences. Shortsightedness and lack of self-command are often related to lack of self-care, as for instance is the case with addiction. In these kinds of activities the negative consequences of behavior manifest themselves only after a longer period. Criminal behavior is also associated with lack of self-command, i.e., an inability to anticipate the negative consequences of one's behavior for the victims and ultimately for oneself. By ignoring the long-term consequences of their actions people restrict the options they need to consider and surrender to their present aims. Awareness of the long-term consequences of their actions requires that the agent deliberates carefully before taking a decision and such a decision can be considered to be motivated by higher-order preferences (see the chapters XVIII and XIX).

20 Contrary to what I suggest here, commitment is in Sen's usage not restricted to moral commitment. Besides commitment based on moral principles people can act on, for instance, sectarian principles. I will often use the concept of commitment as if it refers to moral concerns. (see chapter XIX)
immediately concerns himself, than in what concerns any other man. As a result, man's concern for himself exceeds his concern for other people. Therefore, Smith asked what it is that enables individuals to reach moral decisions. He thought it had to be the common feeling of sympathy. This is the power of the imagination, the capacity of placing oneself in the position of another person and to see oneself from the viewpoint of that person. We should endeavor to examine our own conduct, Smith said, as we imagine any fair and impartial spectator would examine it. Through the principle of sympathy one is able to adjust his feelings and actions to a level that is socially acceptable.

In *The Wealth of Nations* Smith followed another avenue. He wrote that the division of labor, the main factor contributing to a nation's welfare, makes for the increasing interdependence of people. The division of labor is fostered by a desire to serve one's self-interest but in such a way as to engage the self-interests of others. Everyone pursuing his self-interest leads unintentionally, as if led by an invisible hand, to prosperity for all. The mechanism of the invisible hand does not refer to sympathy as an underlying sentiment. Some theorists argued that Smith's portrayal of mankind in the two studies was rather different and that the 'benevolent agent' in *The Theory of Moral Sentiments* (TMS) was replaced by the 'rational egoist' in *The Wealth of Nations* (WN). Some eighteenth-century German scholars of the German Historical School labeled the contradiction inherent in this dual image of man as "Das Adam Smith Problem". In the last quarter of the twentieth century “Das Adam Smith Problem” has gone through three stages. (Montes, 2003) Initially it was labeled as a pseudo-problem, based on ignorance and misunderstanding. It is argued that there are mainly two interpretations of self-interest in Smith's writings: selfishness and prudence. By adopting this second interpretation of self-interest, combined with the essential virtue of justice, it could be argued that Smith's handling of the virtues of justice and prudence just has facilitated the moral acceptance of acquisitive action. (Myers, 1983; Pack, 1997) Thereafter, the
attitude among scholars became that the problem should not be overlooked. The doctrine in TMS may be seen as providing an ethical defense for the commercial society studied in WN. But the price for this solution of "Das Adam Smith Problem" is that whereas self-interest is seen as a motivating force, sympathy was presumed to be a regulative force. This is a disturbing thought and it conflicts with Smith own opinion, for "in TMS sympathy is seen inter-changeable as judgment and as motivation". (Khalil, 1990, 256)

The third stage presents the opinion that "Das Adam Smith Problem" is not solved at all. Witztum, for instance, notices ambivalence in Smith’s writings towards the way in which self-interested people are concerned with the fortunes of other people. Does it come to no more than that they do not intend to treat each other unfair because they have an interest in promoting their own rank and reputation? Or is it a statement that the happiness of other people really matters to them? (Witztum, 1998) Binmore supports Witztum's reference towards these ambiguities in Smith's writings. He believes that sympathy is confused with empathy. He argues that we must look to the distinction between empathy and sympathy in order to achieve reconciliation between Smith’s major books. Sympathy is the attitude that expresses that the well-being of others appears as an argument in one’s own utility function. Empathetic preferences refer to the valuations of the preferences of other agents from the point of view of an impartial spectator. "[T]he solution to the Adam Smith Problem requires reinterpreting many of his references to sympathy in Moral Sentiments as statements about empathetic preferences." (Binmore, 1998, 369). Sugden, however, believes that neither sympathy nor empathy captures the meaning of moral sentiments, and Binmore, as so many

21Another confusion is the equation of sympathy and benevolence. Sympathy means that an agent is concerned with the fortunes of ‘named others’; their well-being directly influences his own well-being. In the case of benevolence the concern for the well-being of other persons may extend to ‘anonymous others’ as well.
representatives of rational choice theory, misrepresents Smith’s intentions. Binmore assumes that we need to recognize that Smith’s definition of sympathy is similar to the modern definition of empathy. 'But sympathy and empathy were brought into rational choice theory for two distinct problems: to explain non-selfish behaviour and to make sense of interpersonal comparisons of utility'. (see Sugden, 2002, 66) Smith’s concern was quite different. Sugden thinks that the term “fellow feeling” best expresses what Smith had in mind when he talked about interdependencies of feelings. For Smith, the psychology of fellow feeling and the correspondence of sentiments are tightly linked with that of approval and disapproval; and approval and disapproval form the basis of the sense of morality. The ideal standard of moral sentiment is to be found in the judgments of the impartial spectator. The psychological mechanisms of approval and disapproval tend to induce norms of propriety of sentiment within any group of interacting people. “...[I]f Smith is right, our sense of what is worthwhile is in part founded on and maintained by the perception of other people’s approval: our consciousness of the correspondence of our sentiments with those of others helps to maintain the sense that our goals are worth pursuing.” (Sugden, 2002, 81)

Though rational agents are pictured as non-tuists, they are not without other-regarding attitudes for they are sensible to the opinion of relevant other agents in their community. They

22 Hume thought that empathetic identification consists solely in the simple imaginary change of position with others, while sympathy implies concern for their welfare. Therefore, sympathy will enhance the possibilities of coordination and cooperation. Fellow feeling could be regarded as linking the notions of empathy and sympathy. This would solve "Das Adam Smith Problem".

23 I have to add that Sugden himself thinks that his interpretation of Smith theory of moral sentiments renders moral sentiments incompatible with rational choice theory, because they are about (affective) states not about preferences. But it seems to me that rational choice theory can as well absorb virtuous persons (whose preferences will by definition be virtuous) as virtuous actions that reveal moral sentiments. (see also the comment by Khalil).
want to continue to trade and exchange with each other and they care to behave in a way that is applauded and feel it as a shame when their behavior is judged as not bona fide.

6. Conclusion

Whereas Marshall defined economics as the study of material pursuits, after Robbins’s influential essay economic science turned into a branch of practical reasoning and ultimately into decision theory. Whereas economics once had a restricted domain, the domain of production and consumption regulated by markets, (micro-) economics afterwards studied an aspect of all human behavior. The change in the object of economics and the formalization of economic theory reinforced each other. In this process key concepts underwent significant changes. First, the concept of value in use was replaced by the concept of utility. Utility refers to a general notion of usefulness. The natural constituents of things were now put out of consideration. (see Meikle, 2001) In due course, the meaning of utility in utility theory has changed. While the principle of utility was once used to indicate that actions are to be judged by their usefulness to produce benefit, happiness, advantage or good, in the modern version a utility function is devoid of any psychological content; it is only a metaphor for the ability of agents to rank their preferences consistently. A utility function is construed as a choice indicator. As a complement to the transformation of the utility concept a logical concept of rationality was adopted. Rationality became an equivalent of consistency (ignoring that one can be consistently malevolent). Rationality so defined might implicate that rationality is no longer associated with self-interest. But prudential concerns are self-regarding and economic man is still constructed as a selfish human being.

The attempts to eliminate all mental concepts from rational choice theory, reduces its explanatory power, for there must be grounds for a choice. As a consequence of the elimination of inner processes and the transformation of the utility
function into a choice function, there has been a behaviorist reconstruction of rational choice theory. Behaviorism eschews recourse to hedonism in its explanatory framework. Rather, behavior is explained by referring to the process of ‘operant conditioning’. In such a process “pleasure” is transformed from a teleological motivating role to a mechanical motivating goal as a reinforcing consequence of past behavior. The “stimulus-response reinforcement” theory is only reinforcing “responses” consequent upon outcomes of past behavior which functions as a motivator of present action. But we cannot deduce from responses in the past the content of the present and future ends and objectives and this is what we must appeal to in order to explain purposive action.

Furthermore, to restore a link between choice and satisfaction, we must include some psychological factors. Thus, as soon as choice is related to well-being, the concepts of utility and rationality must regain some of the psychological meanings they had lost before. Rational action is more than consistent action (see chapter III).

When proposing to include psychological meanings in rational explanations, I am not suggesting that rational choice theory should embrace some kind of psychological individualism. Psychological individualism has had a rather strong influence in neoclassical economics. The reason for this is that the basis of psychologism is that there is something that all individuals have in common. This common element is called ‘Human Nature’. What constitutes human nature is subject of discussion, and today it is merely asserted that the same 'laws' of psychology govern all individuals. In its simplest form psychologism would have us believe that any two individuals facing the same situation would behave in exactly the same way. Simple psychologism allows differences between the choices of individuals to be explained only in terms of the nature-given situations facing the two individuals. All individuals are, in effect identical.\textsuperscript{24} Thus

\textsuperscript{24} In the first section I already referred to the fact that economists take
simple psychologism precludes individuality! (Boland, 1982, 34) An alternative to simple psychologism denies the uniformity of human nature and instead claims that there are different types of people. Thus, when two individuals face the same situation but respond differently, one could explain the differences as the result of the two individuals being of different psychological types, or having different mentalities. (Harsanyi endorses both views; see chapters III and IV)

The central feature of psychological individualism is that only individuals can have aims and that aims are considered to be psychological states. Popper rejects the identification of aims and psychological states. In his view human practices are imbued with normative expectations. They embody ideas about what one is entitled to expect of people and are reinforced by guilt and shame in the face of reproach for the failure to live up to them. (see Hollis, 1994) Rational people know when which rules to apply and how to use them. Without using an institutional framework in which actions are situated or embedded, one cannot explain why people obey rules, commit themselves to principles etc. Institutions are to be included among the explanatory variables that define the aims of individuals. "Behavioral economics" is the answer to the "escape from psychology." (Giocoli, 2003, 41ff) It is an attempt to introduce in economic theory some of the theoretical and methodological approaches of psychology and it introduces rule following behavior. It is a focus on actual behavior, as opposed to the normative ideal of rational behavior. 25

preferences as given. Changes in choice-behavior are explained by changes in opportunity costs. When for example the utility function of a person depends on two commodities, X and Y, and the utility function, therefore, is U: U(X, Y), and the marginal costs of X are declining while the marginal costs of Y remain the same, then the prediction is that this person will consume more X at the expense of Y, assuming that the marginal costs of Y remained unchanged. The advantage of this approach is that we need not refer to (unobservable) changes in the person’s tastes and preferences, all we need to do is to assume changes in the available resources. However, this argument requires that the person’s preferences among X and Y remain the same.

25 It is unfortunate that the label 'behavioral economics' has been chosen,
To conclude I want to say a few things about rational choice and self-interest. Opposing the self-regard versus other-regard distinction I have referred to Parfit's distinction between 'short-sightedness' (present aims) versus 'far-sightedness' (self care over one's life time). I think that Parfit's distinction has two merits. First, it disentangles the concept of rationality and the concept of self-interest. Second, it introduces lower-order and higher-order preferences. The advantage of this approach to rational choice is that instead of combining rational behavior with self-regard, it places rational behavior in the context of preferences with qualitatively different contributions to the well-being of agents. I propose to adopt Parfit's distinction in a modified version of rational choice theory (see chapter XIX and XX).

In the next chapter I discuss some methodological aspects of rational choice theory that I coin as the Model of Desires and Beliefs (MDB). The reason to use this term is that there happens to be an alternative approach, Popper's Model of Situational Logic (MSL). The question is whether these approaches are the same, only appearing under different names, or are competing alternatives. I shall defend the view that they are different, though they both belong to the family of rational choice models. The MDB explains actions in terms of reasons for actions, while the MSL explains actions in terms of constraints on actions.

---

because of its associations with 'behaviorism' that it just wants to replace. Behavioral economics is methodologically the opposite of behaviorism. Because of the possibility of confusion I prefer to speak about institutional or social economics instead of behavioral economics.