Rationality and irrationality in understanding human behaviour. An evaluation of the methodological consequences of conceptualising irrationality

Cosmin Toth[1]

Abstract
Some of the most known and fertile models for understanding human behaviour are those which rest on the assumption of human rationality. These models have specific strategies for dealing with situations in which understanding human behaviour becomes difficult, i.e. cases of irrationality, and this, in turn, leads to particular methodological consequences. The aim of this article is to illustrate and systematize some of the typical theoretical approaches to the issues of rationality and irrationality and their methodological consequences, while warning, at the same time, against the risks of applying rationality models of a pronounced normative-evaluative nature. A number of important methodological consequences of applying the principle of charity to various degrees of strength are analysed and a taxonomic grid for the different ways of approaching rationality is presented.

Keywords
Rationality, rational action, irrationality, principle of charity

Introduction

Irrationality is often invoked when we are unable to understand a form of human behaviour. This misunderstanding can be said to be total, meaning that a particular action or belief is completely unintelligible, or, it could be only the consequence of applying a particular human model to some actions and beliefs in various social sciences. The second case can be noticed when the researcher attempts to find and apply a model which is as simple as possible, coherent, and has a high explanatory power. This second

---

[1] Faculty of Sociology and Social Work, University of Bucharest, Romania, paulcosmin@yahoo.com

case makes the main focus of this article and rests on the assumption that an action is understood when it is translatable in the terms of the adopted model.

Different models which rest explicitly on the notion of rationality as a tool for understanding and explaining human action vary significantly in their approach and classification of human actions and beliefs. The variation is strongly connected to the way in which these models approach and manage that which seems unintelligible, and consequently, to the way in which irrationality is defined. The goal of this paper is to outline some of the significant ways in which irrationality has been approached by rationalist theories and to evaluate the methodological and epistemological consequences involved. Given that it's not possible to review a large amount of works on the topic of rationality, I have chosen to discuss only a number of approaches to irrationality which I find to be significant, and to illustrate each of them by presenting the perspective of one or more relevant authors. This article is not aimed at summing up the existing multitude of rationality models, nor is it intended to argue for a particular model, but instead it means to offer a simplified typology designed to clarify the methodological consequences connected to explanatory models resting on the assumption of rationality. The notion of rationality is often invested with explanatory and normative power, which, if left unexplained, generates confusion. I hope that the typology and examples provided by this paper will enable easier identification of the risks and methodological consequences involved by using or implicitly assuming certain types of rationality models within the explanatory process.

I will start with a short overview of the theoretical and ideological grounds invoked for using the concept of rationality in the social sciences for explaining human action. Next, I turn to the issue of the diversity of models centered on the rationality notion and examine the types of challenges they face when confronted with realities which cannot be easily explained, i.e. cases of irrationality. I start from the assumption that the way in which irrationality is defined is an essential component of these models, one that has significant methodological implications. The argument and typology I present is inspired by Taghart and Nisbett's discussion regarding the use of the principle of charity in understanding human action. I will argue that the discussion regarding the application of the charity principle in varying degrees of severity must take into account the restrictivity of the rationality standard of the theoretical model. The use to varying degrees of these two principles is linked to specific strategies and methodological reactions concerning irrationality. In what follows, the paper suggests classifying some important rationality models according to the two criteria mentioned above. Applying the charity principle in a severe form is associated with defining rationality-irrationality in methodological terms, while relaxing this principle involves accepting irrationality as a real phenomenon which needs to be studied. Depending on the methodological or ontological definition of rationality/irrationality, relevant theoretical examples are presented and discussed for each category.

Rationality as a premise for understanding

Understanding as an analytical endeavour is an established methodological strategy within the Weberian interpretativist tradition, and it rests on the assumption that rationality makes human actions intelligible. Equating rational with comprehensible in this case becomes an attractive option. For instance, von Mises (1966, 52) linking understanding, as a mental research tool, to the field of history, argued that it involves following an interpretative standpoint towards historical events by looking at the means, ends, and motives of relevant actors (mechanisms of rationality). The procedure also involves identifying the contextual factors and their impact on the actions of individuals, as well as assessing the effects of a certain action and the relevance of motives and actions. This is an endeavour of objectively reconstructing events by using actors' subjective constructs relevant for action.

According to Weberian methodology[2], this interpretative 'treatment' of action is done by taking into account two possible layers: an emotional one, where understanding involves 'empathy' (insignificant, however, from the perspective of the researcher's aims), and a rational one. In this last case, the objectives of understanding are either logical or mathematical sentences, or actions involving identifiable ends and means. Specifically, the interpretative process involves reconstructing individuals' reasons by placing them in a wider context of meaning[3] and revealing, as much as possible, the rational grounds of action. The explanation of social phenomena cannot skip this stage. Weber emphasises that explanation is not complete until the entire methodological process of unravelling the subjective meaning of action, phrased in terms of motives, is completed[4].

Boudon, building on Weberian methodology (2003), notes that opting for explaining social phenomena by starting with the rationality of individual actions has the essential epistemological advantage of removing the 'black boxes' from the explanatory process. He also draws from some of the main proponents of rational choice theory such as G.Becker (1996), M.Hollis (1977), or J.Coleman (1986) and argues that rationality as an explanatory mechanism is self-sufficient. Essentially, the argument rests on the finding that unlike any other explanatory option, once an action is understood in terms of the reasons attached by an individual to action, making any further effort to explain is

---

[2] Developed through a dialogue with misesian methodology.

[3] Unlike understanding through direct observation, Weber's explanatory understanding (1947, 95) involves identifying the motives behind the action as well as the context of meaning. In this way, if through direct observation we understand the logic of 2+2=4, when we write it on a piece of paper, through 'explanatory understanding' we uncover both the motives why this is being written on the paper, at this very moment, as well as the context which enabled these motives to emerge.

[4] Weber (1947) emphasises that in order to speak of a sociological approach, statistical uniformities, phenomena should be analyzed only as manifestations of understandable subjective meanings connected to a course of social action: *Statistical uniformities constitute understandable types of action in the sense of this discussion, and thus constitute sociological generalization, only when they can be regarded as manifestations of the understandable subjective meaning of a course of social action* (Weber, 1947, 101).

redundant. Invoking any other cultural, psychological, or social forces would appear to be lacking in clarity and simplicity, creating the need for further explanation.

Boudon supports his argument by invoking alternative ways of explaining human actions, which generate 'black boxes', by ignoring the rationalist interpretation of the reasons behind actions. The Freudian psychoanalytic approach and the methodological premises developed by Durkheim are suggested as examples of such explanatory systems[5].

A number of external political and ideological aspects seem to also support the salience of explanatory models centered on human rationality[6], giving them a competitive advantage. In daily talk, nothing is more legitimate in justifying a behaviour than presenting the grounds for action in an immediately recognizable rationalist vocabulary (goals, means, interests, costs etc.). Moreover, rejecting the notion of reason or rational human behaviour is almost tantamount to rejecting the idea of individual freedom and accepting the existence of hidden forces which govern individuals' behaviour beyond those motives they *artificially* express. In the broader area of economic liberal theory, the Austrian school (see Menger, 1892; Böhm-Bawerk 1890; Rothbard, 1973; Mises, 1966; Hayek, 1944) has built an important debate, with significant imprint on Weberian actionism, centred on the need to introduce rationality in the understanding of human action. The proponents argued in favor of adopting the notions of rationality and individual freedom in order to gain a better understanding of the mechanisms of the market. The two assumptions of freedom and individual rationality were essential prerequisites for their effort to explain things such as: the value of goods on the market, the balance of the market, the emergence of institutions as an effect of the aggregation of individual decisions and actions, and the functional advantages of these institutions compared to any centralized forms of economic systems.

Under these circumstances, the assumption of rational individuals and rational actions seems to be an approach with methodological advantages and ideological or doctrinal legitimacy. Few other conceptual tools seem to benefit from the advantages of such simplicity, coupled with such a high explanatory power, and backed by strong legitimacy in the space of public discourse.

Rationality norms, the principle of charity and the place of irrationality

We are presented with an image according to which once human actions are translated into a system of rational motivation, this becomes an almost indisputable explanation

---

[5] Boudon argues that the methodological prescriptions promoted by Durkheim are not, in fact, applied in his explations of social phenomena. Durkheim is rather using a weberian methodology of explaining social phenomena by unravelling the grounds of individuals' actions which are behind it.

[6] I have to note that the association between the individual and rationality is a fundamental premise of the analysis presented in this paper. I will only discuss the issue of rationality/irrationality in this sense. According to this approach, rationality makes sense only to the extent that it is linked to the individual who acts, intends, deliberates. I will not approach the notion of rationality used in connection to entities other than the individual, such as systems, ways of organisation/aggregation, or collectivities.

and there is no need for further clarification. However, a number of problems seem to arise precisely when such a translation is intended. Without insisting here on these aspects, it is enough to mention two of the problems of the interpretative process, which Weber himself noted: on one hand, the cultural distance between the researcher and the subject who's motives (in terms of ultimate ends and values) the sociologist is supposed to understand (1947, 91), and on the other hand, the possibility that there is no overlapping between what the subject invokes as motives for acting and what the researcher manages to grasp and interpret as such (1947, 97). A characteristic feature of this translation is the fact that it involves following a rationality model and the problems which arise stems exactly from the key of this *translation.* We can take this key to be the criteria used in order to distinguish rational actions from other actions. The key should offer the definition of rational action, the criteria according to which we distinguish rational actions from other actions, as well as the way we should deal with the latter. This task of differentiation becomes problematic when we are unable to express motives in a meaningful system. In general, by contrasting them with rationality (as intelligible), these actions are defined as irrational. Dealing with those actions, decisions, or beliefs, which do not seem intelligible in terms of the assumed rationality model, becomes a major methodological challenge.

Anthropological research, as well as our everyday experience, presents us with examples of people, decisions, actions, or beliefs which are capable of provoking a variety of reactions, ranging from a simple puzzlement, to the ascribing of labels such as: irrational, illogic, primitive, unintelligible, or even worst, uneducated or mad. For instance, in western society, alternative belief systems regarding the workings of the world, such as astrology, magic or religion, together with their followers and practices, are often the target of such labeling.

However, beyond an initial reaction in the realm of common sense which may not even spare social researchers, there is an epistemic problem which I will focus on in what follows.

Taghart and Nisbett (1983) examine whether the principle of charity[7] can be applied in three areas where understanding is involved: translation, inferential processes and decision. The principle of charity essentially recommends that the process of understanding any inference or decision should rest on the principle of the logical non-contradiction and rationality of the source. As a consequence, the recommendation is that whenever the researcher is unable to understand, he/she should avoid attributing the responsibility of the misunderstanding to the source.

Thagart and Nisbett (1983) suggest that the principle of charity can be conceived as having varying degrees of strictness, ranging from banning the interpretation of any action/decision as being irrational, to the absence of any prior recommendation regarding rationality or irrationality. The authors argue for a medium level of severity, expressing the following advice: 'Do not judge people to be irrational unless you have an

---

[7] Concept taken by the authros from W.V.O. Quine.

empirically justified account of what they are doing when they violate normative standards.'(1983. 252)

The authors build their argument on the assumption that there is a given, generally accepted normative standard used in distinguishing between rational and irrational actions. If this is generally clear where logic is concerned, things are not so simple in the case of evaluating a decision, a belief, or an action. For the latter situation, the authors seem to be guided by the expected utility theory, to the extent that they challenge the principle of strong charity by using Tversky's and Kahneman's experiments. Although it is true that this is the most known standard for defining rationality, there are other ways of defining and understanding this concept, which apply different standards of evaluation. While the theory of expected utility (as a formal model of rational choice theory) implies a severe, restrictive standard of rationality, there are other models, more flexible and relaxed, which identify as rational even those actions which are carried out because of inertia, tradition or habit (Boudon, 1997). I will argue in what follows that when discussing the application of the principle of charity we must take into account what degree of strictness of the criteria/standard of rationality has been used and I will also try to describe some of the methodological consequences generated by using the two principles to varying degrees of strictness.

While in the case of a restrictive model of rationality, the advantages noted by the authors quoted above concerning a charity principle of a medium severity can be maintained, as the rationality criteria become more lax, the principle of charity can become more severe without raising the aforementioned problems. It is obvious that adopting a high level of severity for both standards bears the promise of incontestable epistemic advantages: the severe principle of charity is an assurance that the explanation and understanding of human action will have no black boxes, while formulating restrictive rationality criteria allows for articulating a unified theory of human action, phrased in a small number of axioms, which are easy to apply. This type of undertaking, discussed and amended by Thagart and Nisbett (1983) and successfully applied in sociology by J. Coleman (1990), generates methodological problems, such as the forced explanation of beliefs and actions, which are difficult to fit within an instrumentalist, utilitarian model of action[8].

On the other hand, there is the option of choosing to maintain a severe principle of charity at any cost, accepting at the same time the possibility of flexible rationality standards. This is the case with the methodology suggested by Boudon (1997) who deems primary the exclusion of black boxes from the explanation, enhancing the model with a multitude of forms of rationality, together with instrumental rationality (perhaps the most restrictive), such as axiological rationality, traditional rationality, teleological rationality, cognitive rationality, and so on, leaving the list open (relaxing, in this way, the rationality norms).

---

[8] See the example given by Thagart and Nisbett (1983, 262) who explain non-optimal decisions taken by gamblers by the fact that they are maximising non-monetary gains such as self-esteem or amusement.

Equally, it is possible to opt for maintaining severe rationality criteria, while at the same time, accepting the possibility that irrational actions or decisions exist. This is the methodological stance promoted by Thagart and Nisbett (1983) and successfully employed by J. Elster (1983). In this last case, once the charity principle is relaxed, a significant part of the explanatory endeavor moves towards understanding and explaining irrational actions. Together with the abandonment of the severe principle of charity, comes the need to accept irrational phenomena, which demand new models of understanding and explaining human action.

In this way, as the charity principle becomes more relaxed, the researcher's effort moves from the methodological refining and adjusting of the model to new situations, towards the ontological acceptance of new categories of phenomena which demand new research models.

Figure 1. Graphic outlining the possible positioning of different theoretical models according to the degree of severity of their principle of charity and to the restrictiveness of their rationality norms[9]
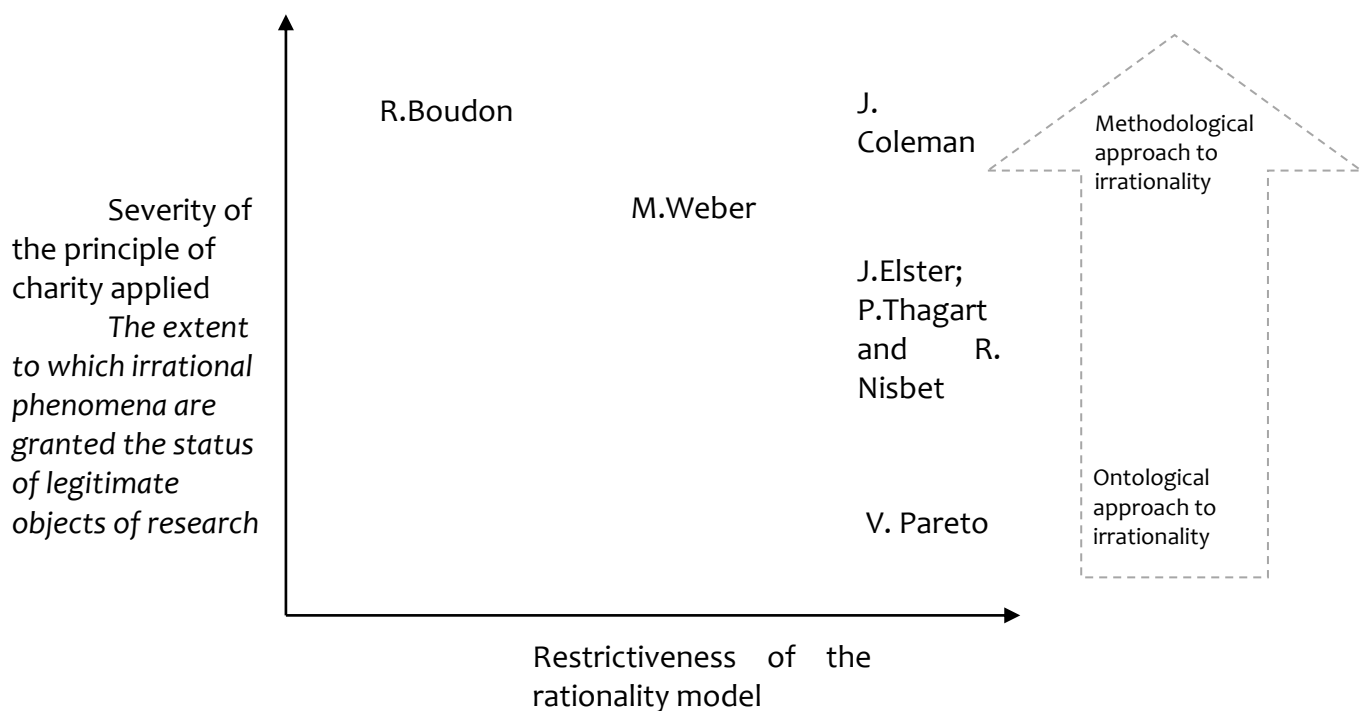
Consequently, variation in the severity of the principle of charity generates different approaches, particularly in those borderline situations where irrationality

---

[9] I have only sketched this graphic in order to offer a summarising and simplified visual representation of the way in which certain rationality models could be positioned according to the principle of charity they apply and to the more or less restrictive rationality standard they use.

ensues and difficulties arise in the understanding of human behaviour, decisions, and beliefs. Next, I will describe a number of strategies connected to the acceptance or rejection of irrationality as an object of research, stemming from the use of the principle of charity.

The first option in these situations (irrationality) is to adopt a methodological strategy based on a severe principle of charity. Discovering the model's incapacity of translating the action or belief can trigger the following types of answers:

- The model might be wrong. In this case, a procedure of improving/enriching the model can be initiated in order to include the new types of actions identified.
- Information might be insufficient. The obvious answer under this circumstance points towards further investigation. Misunderstanding stems from the lack of information or even from cultural limitations. As a convention, the action will continue to be seen as irrational, albeit only temporary.
- This type of action is atypical, rare. In this situation, the effort of adjusting the model to a limited number of cases seems unjustified. The existence of such actions can be admitted, but they are considered unimportant for research.

A second alternative can be seen as an ontological approach determined by taking into account the possibility that irrational phenomena exist, through a relaxing of the principle of charity:

- There are people/actions which are rational and actions which are irrational. There are clear mechanisms/criteria for distinguishing rational actions from irrational ones. Processing, classifying and analysing these situations of irrationality becomes an important strategy. In this case, the line of intervention can either go towards trying to modify the reasoning process of the person/category, with the risk of turning into a socio-centric approach, or, it can lean towards accepting irrationality as a separate phenomenon, a trait of the human psyche which exists and needs studying in the same way as rationality does.

Both methodological and ontological approaches establish certain standards and imply an evaluation leading to a normative approach: 'If one would have been rational, he/she should have acted in such way, but he/she did not'. What makes the approaches differ is the fact that within the methodological stance, the blame is placed on the model or on the lack of information (methodological normativity), any intervention being aimed at the improving the models, while in the case of the ontological approach (ontological normativity), the person and his attributes become blamable.

In what follows, in order to illustrate the consequences generated by various approaches to irrationality, I will carry out an analysis focused on a number of examples of theories which take into account the issue of irrationality/rationality of human actions and beliefs.

Rationality as a methodological tool or as an ontological premise

According to the classification presented above, I will group different approaches to irrationality into two categories, according to the way in which they define the notion of rationality, and depending on their intended type of evaluation of human actions, decisions, and beliefs. For each of the approaches that I will describe and analyse, their particular approach of irrationality has important consequences regarding the way in which research is designed and guided, and later on the way in which the explanation is built. Ultimately, the way in which irrationality is approached specifies what makes a legitimate research subject, as well as the lines along which the explanation is built, together with the limits of the research and the ways of dealing with these limitations.

The first identified category is that in which the notion of rationality is used as a methodological tool constrained by a severe principle of charity. The second category is that which rests on a relaxed principle of charity and follows a number of ontological assumptions concerning the rationality of individuals, regarding irrational actions and beliefs as significant phenomena.

Rationality as a methodological tool

Within this type of approach, no assumptions are made regarding individuals' actual rationality, but instead the explanatory/methodological utility of the concept is explored and exploited. Irrationality is regarded as a 'difficulty' in translating action in terms of a rationality model and dealt with either 1. As a temporary limitation of the range of phenomena which can/should be explained, or 2. As a false problem.

In the first case I will invoke Weber's, Boundon's and Coleman's perspectives on irrationality. Each of them admits the ontological fact that individuals are not entirely rational and that irrationality is constantly present in their actions and manifestations. On the other hand, they recommend conceiving rationality as the instrument with the highest explanatory power and starting with the assumption that individuals are first of all rational (the charity principle). Within these theories, efforts are directed towards explaining as much as possible by using the tool of rationality. Irrationality is accepted as an ontological reality, but its nature is not analysed or theorised, and there is no concern for the need to understand its various manifestations. Irrationality is an area which needs precise limiting and delineation in order to leave as much room as possible for rationality as a chief explanatory tool. Within this group of theories, variations in the usage of the concept of rationality are determined by the optimal ratio between the simplest, most narrow and applicable conceptualization of rationality (restrictive criteria) and the highest possible explanatory power, given by the volume of phenomena that it is able to clarify.

Coleman (1990) takes a ‚utilitarian' stance arguing that individuals should be regarded as rational and as maximizing their satisfaction (utility). He recommends caution in dealing with seemingly irrational acts, as most often they can be explained by rational mechanisms. As an example, he notes the case of apparently irrational

revolutions rising on the backdrop of an apparent improvement in the livelihood of a society previously afflicted by poverty.

> Here, as in other cases, however, a valuable theoretical strategy is to **refuse** to accept the apparent **irrationality** and instead to ask what perspective might make it rational in the way tht people are observed to act. (Coleman, 1990,p.472)

Gary Becker (1962), another proponent of rational choice theory, defines as rational only those behaviours that involve maximising utility. However, he adds that this is only a convention and he defines as irrational two types of behaviour: impulsive and inertial.

> Any deviation from utility maximization is considered irrational in this paper: a more precise or philosophical definition is not requiered for our purposes and is not attempted. (Becker 1962, 5)

Along the same lines, Boudon argues for avoiding and limiting as much as possible those phenomena which are the result of *apparent* irrationality. However, in order to delineate the sphere of irrational phenomena, he uses a much more relaxed, semantic definition of rationality. Boudon considers irrational those behaviours which cannot normally be described with the help of expressions such as 'X had reasonable motives for doing Y, because...' (1997, 47), without creating an image of absurdity. For instance, he notes that a phrasing such as: '*The mother had sound reasons for slapping her child because she was angry.... (1997, 39)* triggers ambivalent feelings. In this case, the contradiction is generated by the attempt to translate reasons of an affective nature into motives that are phrased in deliberative, cognitive terms.

By using a semantic criteria, Boudon (1997) broadens the range of actions classified as rational according to the Weberian definition. While Weber's methodological recommendation (1947, 92) is to approach traditional and affective (emotional) actions as irrational, as deviations from the ideal type, Boudon (1997) tends to include them on the list of rational actions, even leaving the list open for adding other actions , based on semantic tests. He speaks, for instance, of traditional rationality expressed as: 'X had sound reasons to do Y because X has always done Y and he/she had no grounded reason to question this practice' (Boudon 1997, 42). The semantic definition leaves open the issue of the relative nature of the degree of 'soundness of motives'. For instance, it is not clear whether terrorist actions are classified as irrational.

The second situation in which rationality is used as a methodological tool addresses a different epistemic level and disregards the issue of irrationality completely, considering it to be a false problem. This is the case with von Mises's Praxeology and with the Popperian rules. According to these models, the idiom 'rational action' is a tautology. Rationality is a principle which is meant to animate any theoretical model.

Ludwig von Mises (1944) writes from the standpoint of Praxeology arguing that a science of human action should not take into account rationality since it is a quality produced through assessment. Judging and classifying an action as rational or irrational

is not the object of a neutral and objective science aiming to formulate laws of human action a priori. A science of action should not assess peoples' goals, nor should it evaluate the means they choose for achieving those goals.

> Praxeology does not employ the term rational. It deals with purposive behaviour, i.e., human action. The opposite of action is not irrational behaviour, but a reactive response to stimuli on the part of the bodily organs and of the instincts .... If it were to assign a definite meaning to the term rationality as appied to behaviour, we could not fiind another meaning than: the attitude of men intent on bringing about some effects.(Misses,1944,pp.533-534)

> The terms irrational and irrationality are mostly used for ensuring concrete modes of action. An action is called irrational either because the censor **disapproves** of the end or because the censor believs that the means employed were not fit to produce the immediate effect aimed at. (Misses, 1944,p.534)

> As far as there is scarcity of means, man behaves rationally, i.e., he acts. So far there is no room left for "irrationality." (Misses,1944,p.544)

A similar line of reasoning is developed by the philosopher Karl Popper (1998) who argues that rationality is the only valid principle that can animate our explanatory models. This principle does not oppose the falsifiability principle because it is false from the beginning. Popper's argumentation touches on the following points:

1. Explanation and understanding in the social sciences are only achieved through situational analysis. The aim is to develop models of typical social situations.
2. In order to animate these models, we need to assume that people act *appropriate to the situation, sensible.* This is the principle of rationality. This principle has nothing to do with empirical or psychological assertions concerning rationality.
3. This principle does not play the role of a theory or a hypothesis. This principle is essentially false because it is a simplification and an approximation of reality. Natural sciences also use such simplified models.
4. In any explanation there is a situational model animated by the principle of rationality. Should the explanation fail, the recommendation is that the blame should be placed with the model, and not with the principle behind it. It is far more interesting and instructive to improve situational models through trial and error than to try to increase our capacity to act adequately and to understand what is appropriate, especially since albeit the principle of rationality is false, it approximates the truth quite well. Using any other principle only makes a model arbitrary. Everybody acts appropriate to the situation and according to their knowledge, even the madman.

The theoretical core of this type of models is based on the methodological requisite that action is considered to be rational by default. Within this context there is no question whether actions are rational or not and the rationality model/concept is positioned a priori in relation to subsequent constructions, representing rather an axiom used as a starting point for further theorising.

Rationality/irrationality as an ontological problem and premise

When rationality is taken as an ontological premise, the fundamental problem is determining the criteria by which an action, decision or belief can be classified as rational or irrational. Within this group of theories two ways of applying these criteria can be identified.

The manner of defining the criteria for the first category is the most questionable. Generally, this involves attributing an irrational nature to unintelligible actions. I include in this category those arguments where rationality is overlapped, at least to some extent, with notions such as logic, intelligence, or education, and, conversely, irrationality is linked with ideas of missing logic, errors in logic, or even with a lack of intelligence or education.

An illustrative and meaningful example is Pareto's distinction between logical and non-logical actions.

Summing up briefly, for Pareto (1965, 77) a social phenomenon is objective if it is true to reality and it is subjective in the way it is presented to our mind. Logical actions are those actions that employ the right means for attaining certain goals, those actions that connect means to ends through a logical connection which seems logical not only to the person who acts, but to some other person with superior knowledge. For these actions the objective end is the same as the subjective goal. Non-logical actions are those where there is no objective link between means and ends and the objective end differs from the subjective goal. The gap between the subjective assessment of an action's link to its end, and the objective reality, allows for identifying four types of non-logical actions. The resulting non-logical actions are: customs and habits, prejudices, instinctual actions and utopic projects (Pareto, 1965).

Unlike rationalist theories, Pareto suggests that non-logical actions are dominant in society and therefore sociology should focus first of all on analysing and following the manifestation of non-logical actions. According to this, Pareto's distinction seems to overlap with the generic definition of rationality (see Boudon 1997). If we rephrase the definition of logical actions we could assume that those are the only actions Pareto considers rational.

I have invoked Pareto's classification of actions because it presents a clear and distinctive way of approaching the problem of irrationality. What is distinctive for this approach is the fact that rational and irrational actions are differentiated by following the criteria of scientific objectivity (with specific methodological consequences) and a restrictive rationality standard is applied simultaneously with a relaxed principle of charity.

Winch (1990) is among those who construct a critique of the 'objective' nature of the criteria used in assessing rationality. I describe this critique here in order to illustrate the traps of such an undertaking. According to Winch, concepts and discourses are no longer comprehensible once they are artificially extracted from the life setting where they were generated and functioned. Theories about the workings of the world and the actional approach to the world employ concepts which are particular to certain life forms and cannot be extrapolated without losing their meaning.

In this way, Winch (1990) claims that Pareto makes a number of fundamental errors when he classifies human actions in logical and non-logical. The problem lies with the criteria according to which this categorisation is achieved.

> On the other hand, to try to understand magic by reference to the aims and nature of scientific activity, as Pareto does, will necessarily be to misunderstand it. (Winch 1990, 100)

> That criteria of logic are not a direct gift of God, but arise out of, and are only intelligible in the context of, ways of living or modes of social life. (Winch, 100)

Pareto's mistake, according to Winch, is that he argued that science is, in itself, a type of logical behaviour, while religion is non-logical. This means that he granted preferential treatment to one life form (science) contradicting in this way the principle of neutrality and scientific objectivity.

It is erroneous to apply criteria used in ranking various scientific theories according to their level of logic (judging by the logical-experimental criteria) to human life forms, and then label those forms non-logical (obviously, apart from scientific behavior) (Winch, 1990).

Winch's discussion of Pareto's theory clearly highlights the problems and risks involved by setting up ontological criteria in order to classify actions according to their degree of rationality. Most of the risks emphasized are of an epistemic and methodological nature and they demonstrate the kind of errors that might emerge in the classification of human actions. The risks become even more important when these taxonomies are used as a framework for evaluating human groups, societies, or their practices (see, for instance, the case of societies said to be 'primitive' or practices labeled as 'magic') ultimately leading to socio-centric labeling.

C. Jarvie and J. Agassi (1967) look at a similar issue discussing the rationality/irrationality of systems of magic beliefs. Analysing the problem of the relative nature of rationality, the authors make an important distinction regarding the way this notion is used. On the one hand, we can talk about rational action, and on the other hand, we can speak of rational beliefs. This distinction aims at eliminating the normative-evaluative judgment of actions' rationality by placing them in relation to a certain cultural superiority.

When we attribute rationality to a person we can mean either: he acts rationally, or he believes rationally, or both. Let us call the rationality that consists in a person acting rationally weak sense of 'rationality'; and the rationality that consists in a person acting rationally on the basis of rationally held beliefs the strong sense of 'rationality'. (C. Jarvie and J Agassi, 1967)

Next I will recall a number of confusions created by authors such as J. Frazer and J. Beattie by disregarding the distinction noted above. The most serious consequence of this misunderstanding is to regard magic-using cultures as primitive, leading to an evolutionist approach of magic – magic as a primitive science. Hence, the authors argue, there is no doubt that those cultures which practice magic are not irrational, but they have a weak form of rationality postulating links between causes and effects. When analysing the means employed by magic-using cultures in order to attain a goal (such as having a good crop) one should not make false distinctions, as for example between magical means and purely technological means. They should be taken as a unit, as one single means:

The problem might become something more like 'is primitive society rational in the weak sense?', i.e. can the actions of people in primitive societies be rationally explained by means of their aims, beliefs and knowledge of their circumstances? There is no doubt about the answer. If we press the point, though, and ask whether primitive people adopt a rational or critical attitude to their beliefs and explanations, the issue becomes a little more sensitive. The answer depends upon one's criterion of rationality; and then it becomes a factual question: does or does not such and such a society possess, e.g. the tradition of adopting a critical attitude towards beliefs, values, explanations? This is a question of sociological fact: the presence or absence of a tradition. It reflects in no way upon the intelligence, stupidity, human dignity, or mental capacities of the peoples of the societies in question. (C. Jarvie and J Agassi, 1967, 69-70)

I have presented Jarvie and Agassi's discussion in order to briefly illustrate how approaching rationality in ontological, evaluative terms can generate problems of cultural categorising.

Beyond a simple reiteration of the problems raised by using criteria of rationality across cultures and the corresponding sociocentric risks, Jarvie and Agassi present us with a solution. They suggest two ways of approaching magic beliefs (in an evaluative manner). If the evaluation is directed at the content of the beliefs (means and objectives) they need to be accepted as rational based on the premise of cultural relativism. At the same time what can be assessed is the believers' approach to their own beliefs. In this case the critical standpoint becomes the criteria for differentiating and for granting the 'rational' attribute.

The solution provided by the two authors continues to bear some of the problems previously mentioned. This becomes visible if we review their questions concerning the critical judgment of people who hold 'inefficient magic beliefs'. The legitimacy of this

question can be problematic as it still carries nuances that can be classified as sociocentric.

Apart from the consequences of using rationality standards typical of the western culture in order to evaluate other cultures, assessments of human rationality can be directed to categories of people within one's own [western] culture. The results of such valuations can even lead to recommendations regarding the social policies targeting particular social groups.

Sandra J. Peart (2000) gives an overview of the ways in which neoclassic economists approached irrationality at the end of the XIXth century (S. Jevons, A. Marshall, I. Fisher, A. Pigou). In all these cases irrationality is regarded as an ontological fact which needs to be researched, but at the same time it is seen as a curiosity.

Peart's (2000) article *Irrationality and intertemporal choice in early neoclassical thought* identifies systematic attempts from the part of the above-mentioned neoclassic economists at attributing irrationality to the working classes by overlapping it with the lack of education and ignorance. People belonging to the working classes are regarded as having a strong inclination for present, immediate income, and for consumption at the expense of saving or postponing consumption, even when delayed consumption is more promising. This is put on the account of variables such as: lack of patience, lack of control, lack of will, little education (the ability to asses correctly long term consequences).

This is how Peart sums up the economists' perspective:

> Yet what is clear is that education was regarded as a remedy for both of these problems. By teaching the poor to look to the future and to plan for fluctuations in labour demand, education would improve foresight; and by inculcating moral values, such as prudence and restraint, education would enhance their ability to resist the impulse to stop by the local grogshop on payday. (Peart 2000, 188)

The cases discussed so far present us with enough grounds for arguing in favour of caution when using models of rationality that may involve normative standards employed to classify people or entire cultures as irrational. Nonetheless, when Thagart and Nisbett (1984) recommend using a relaxed principle of charity, although they do admit a number of ethnocentric pitfalls, they advise to take into account the advantages of keeping aware of individuals' cognitive limitations, precisely for the purpose of improving them. It remains however unclear what are the standards and who sets them, according to which diverse social categories or social practices should they be evaluated in order to initiate an educational policy. Although the authors draw from A. Tversky's and D. Kahneman's experiments and show that people are predisposed to systematic errors of reasoning, considering the option of an educational policy focused on this seems rather unrealistic and even risky because of its sociocentric inclination.

One final category that I wish to analyse is that of those theories which aim at identifying universal objective criteria for distinguishing between rational and irrational actions, conducting a detailed analysis of the resulting cases of irrationality, in order to understand the mechanisms behind them. This line of reasoning tries to overcome the

tautology inherent in the idiom 'rational action' by assuming that generally there are some 'gaps' in the functioning of the human psyche. In this way, an acceptable level of universality for rationality criteria is maintained and sociocentric risks are overcome. As opposed to the previous approach, irrationality here is no longer attributed to people, cultures or beliefs, but it is regarded as a phenomenon which, given the structure of the human psyche, might emerge in the actions of any individual. For this set of theories I took into account in my analysis authors such as J. Elster, D. Davidson, J. Church. The starting point here is an analysis of the likelihood that action is based on deliberative and intentional processes, albeit irrational ones. The irrationality criteria explored by these theories are consistency, constancy, and the degree of coordination between beliefs, wishes and actions. The authors discuss irrationality cases such as *wishful-thinking, self-deception, weakness of will (akrasia)*. Irrationality is approached and analised as a mental phenomenon as present as rationality is.

Donald Davidson (1990) starts from the notion that irrationality is an ill-functioning process or mental state and raises the following question: how can we accept the possibility of irrational acts when by explaining an action through individuals' attitudes, wishes, or beliefs, then his/her action becomes intelligible, reasonable, and hence rational?

Irrationality emerges when, for the same person, there is a dysfunction in the coherence or consistency of the pattern of beliefs, attitudes, emotions, intentions, and actions. Examples are processes such as wishful thinking and self-deception:

> A person is irrational if he is not open to reason – if, on accepting a belief or attitude on the basis on which he ought to make accomodating changes in his other beliefs, desires or intentions, he fails to make those changes. He has a reason which does not cause what it is a sufficient reason for. (Davidson, 1990,458).

In the process of *wishful thinking*, a strong desire becomes a cause for a belief, but not a reason for it. Although the relation is causal, it is not logical. This is a case of irrationality.

To conclude, D. Davidson (1990, 462-463) emphasizes that in order to understand irrationality we need to accept three assumptions (derived from Psychoanalysis): the mind is a partitioned construct, each part has its own structure, and there are causal non-logical relations between the parts.

Starting with the assumption that individuals can experience unreflective intentional states, given that it is impossible for these states to be simultaneously present in a person's reflective consciousness, J. Church (1987, 362) argues that beliefs, desires, and irrational actions are those that have a weak coordination with reflective states and reflective forms of reasoning.

In the same way, J. Elster (1983) describes the irrationality of acts such as the attempt to overcome insomnia by using volitional resources (the insomniac making an effort of will in order to fall asleep) or the careful planning of spontaneous behaviour. These intentions are self-defeating. Temporal preferences can be the object of two kinds of irrationality: impatience and inconsistency (consistency implies that consumption that
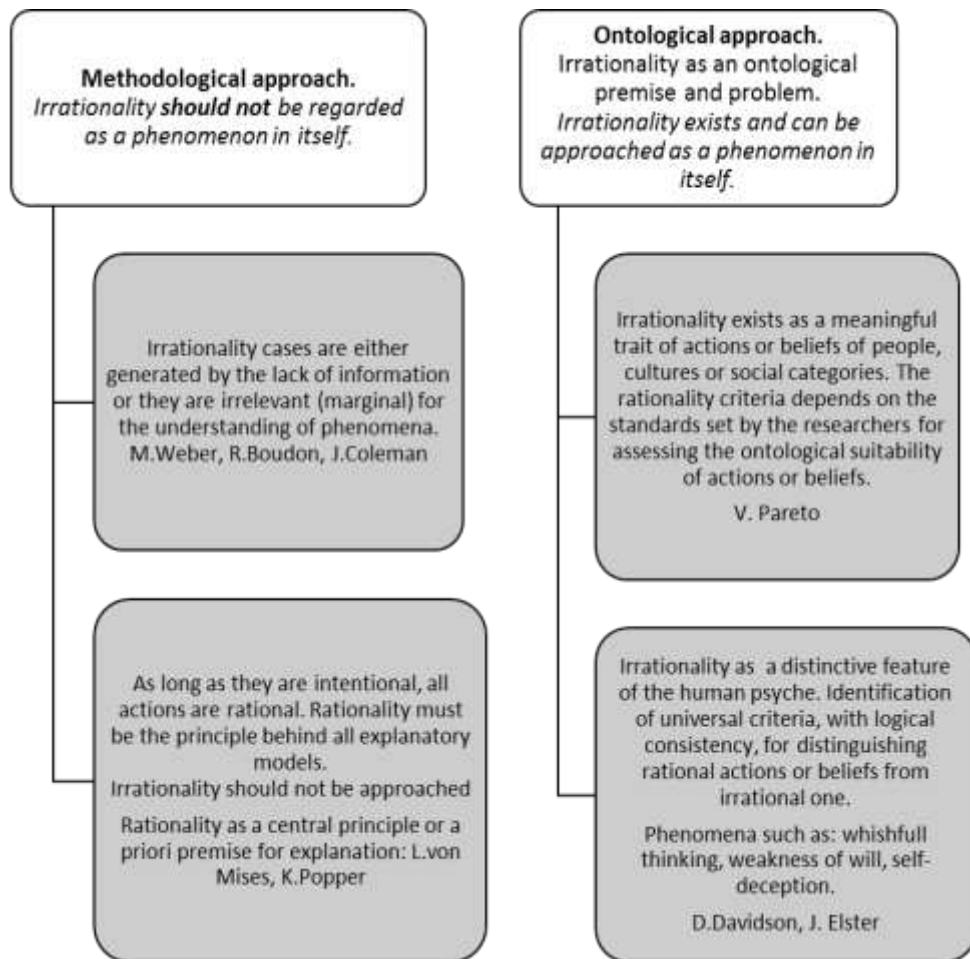
was planned at the time t1 for the t2-t3 period will be still valid at t2 time, provided that there were no changes in the personality or in the feasibility of the set).

In all of these cases, irrationality is analytically assessed in order to establish, on the one hand, the objective criteria for its identification, and on the other hand, its mechanisms and occurrence in individuals' actions, desires and beliefs.

To conclude, a researchers' approach to the problem of irrationality will influence the way he/she represents and deals with the phenomena under study, as well as his/her construction of conceptual tools. If irrationality is understood as an obscure area which needs to be reduced as much as possible, then the research effort will be directed towards refining the concept of rationality (for instance, by extending the definition to a wider array of situations) and phenomena that cannot be understood will be temporary 'marginalised', while hoping that by collecting more information, it will be possible to find a rational explanation for them. On the other hand, accepting irrationality as one of the mechanisms of the human mind (ontological premise) leads to an effort of understanding and describing this phenomena as a specific object of study. In this case, the effort is directed towards establishing scientific standards for delineating and classifying a mechanism as irrational in order to analyse it with precision.

In order to give a summarising, structured image of the argument presented above, I have illustrated through the following graphic:

Graphic 2: Summarizing outline of the different approaches to irrationality

Methodological approach.
Irrationality **should not** *be regarded as a phenomenon in itself.*

Ontological approach.
Irrationality as an ontological premise and problem.
Irrationality exists and can be approached as a phenomenon in itself.

Irrationality cases are either generated by the lack of information or they are irrelevant (marginal) for the understanding of phenomena.
M.Weber, R.Boudon, J.Coleman

Irrationality exists as a meaningful trait of actions or beliefs of people, cultures or social categories. The rationality criteria depends on the standards set by the researchers for assessing the ontological suitability of actions or beliefs.

V. Pareto

As long as they are intentional, all actions are rational. Rationality must be the principle behind all explanatory models.
Irrationality should not be approached

Rationality as a central principle or a priori premise for explanation: L.von Mises, K.Popper

Irrationality as a distinctive feature of the human psyche. Identification of universal criteria, with logical consistency, for distinguishing rational actions or beliefs from irrational one.

Phenomena such as: whishfull thinking, weakness of will, self-deception.

D.Davidson, J. Elster

To end with, I would like to note that my review and theoretical analysis have serious limitations concerning the paradigmatic range they cover, as well as the high level of complexity associated with the scholarship of the theoreticians they draw from. Grasping the entire theoretical complexity surrounding the concept of rationality is far beyond the aspirations of this paper. Without a doubt, important theoretical contributions focusing on the concept of rationality have been omitted and the analytical models that were mentioned were simplified. For the purpose of this paper, the theoretical models that were analysed were employed rather as simplified but significant prototypes for the typology and the discussion presented. I do hope, however, that despite not having a complex and overarching approach, my paper succeeded, through its schematic, simplified models and through the theoretical examples provided, to summarise and integrate in a meaningful taxonomical scheme a few significant

approaches to the problem of irrationality, while at the same time drawing attention to their methodological consequences.

Conclusions

I have started this paper with a short description of the way in which the notion of rationality became the central assumption for the understanding of human action, and subsequently I have pointed out the epistemological advantages and the ideological legitimacy that this concept has for western society. I then invoked as a starting point for the discussion the principle of charity, according to which the subjects of research should be first of all approached as rational individuals, something that bears significant implications and consequences on both epistemological and ideological levels (a principle which is fundamental for the equal treatment of people – see Thagart and Nisbett (1983)). Next, I have illustrated a number of difficulties and methodological muddles that come along with the variability of the degree of severity in the application of the principle of charity, at the same time with the varying degree of strictness of different rationality models. Different ways of applying these principles have different consequences for the research process. Depending on the degree of strictness of the principle of charity and on the restricting nature the of the chosen rationality model, the researcher will focus either on the methodological refining of the models (rationality as an instrument/methodological principle), or on recognising that irrational phenomena can be a research area (ontological approach). For each type of approach I have identified two subtypes and I provided examples of relevant theoretical standpoints. On a methodological level, irrationality is seen either as a temporary misunderstanding, insignificant for the sociologist, or as a false problem. From an ontological point of view, irrationality emerges as an object of research, generated either by applying more or less arbitrary standards of rationality to certain categories of people, or through identifying processes of logical inconsistency in the human psyche.

A lack of comprehension or insufficient details concerning the ways in which the notion of rationality is employed generates important risks. Usage of the concept is prone to lead to evaluations. As the notion of rationality is 'contaminated' by western ideology, it can easily generate unacceptable evaluations, determined by arbitrary rationalist standards. Consequently, one of the goals of this paper was to draw attention to the difficulties involved by using the concept of rationality in social research. Although rationality seems a promising conceptual tool for understanding human actions, decisions and beliefs, it also requires rigorous epistemic delineations.

R E F E R E N C E S

Agassi J., and Jarvie, I.C. (1973). Magic and Rationality Again. *The British Journal of Sociology*, 24, 236-245

Becker, G. (1962). Irrational Behavior and Economic Theory. *Journal of Political Economy, 70*, 1-13.

Boudon, R. (1997). Acțiunea în R. Boudon (coord.), *Tratat de sociologie.* București: Humanitas.

Boudon, R. (2003). Beyond Rational Choice Theory. *Annual Review of Sociology*. 29. 1-21

Bohm-Bawerk, E. (1890). *Capital and Interest A Critical History of Economical Theory.* London: Macmillan and Co.

Coleman, J. (1990). *Foundations of Social Theory.* Cambridge: Harvard University Press.

Church, J. (1987). Reasonable Irrationality. *Mind, 96*, 354-366.

Davidson, D. în Moser, P. coord. (1990). *Rationality in Action*, cap. "Paradoxes of Irrationality", Cambridge University Press

Elster, J. (1983). *Sour Grapes.* Cambridge: Cambridge University Press.

Hayek, F. (1944). Scientism and the Study of Society. *Economica, 41.*

Menger, C. (1892). On the Origin of Money. *The Economic Journal, 2*

Pareto, V. (1965). *A Treatise on General Sociology.* New York: Dover Publications.

Peart, S. (2000). Irrationality an Intertemporal Choice in Early Neoclassical Thought. *The Canadian Journals of Economics, 33.*

Popper, K. (1998). *Mitul contextului.* București: Trei.

Rothbard, M. (1973, trad [2001].). *Ludwig von Mises pe înțelesul tuturor*, available on *http://mises.ro/244/.* Accessed on 10.09.2012, www.misesromania.ro, Institutul Ludwig von Mises – România

Thagard, P. și Nisbett, R. (1983). Rationality and Charity. *Philosophy of Science, 50*, 250-267

Weber, M. (1947). *The Theory of Social and Economic Organization.* New York: Oxford University Press.

Winch, P. (1990). *The Idea of Social Science*. London: Routledge

Cosmin Toth, University Assistant, Faculty of Sociology and Social Work, University of Bucharest, PhD. with the thesis: "Rationality Models in Social Sciences", University of Bucharest. Teaching the following courses, seminars and laboratory work: Sociology of Organizations, Management, Organizational Behavior, Decision and Decision Techniques, Organizational Conflict, Organizational and Administrative Culture in Romania, Decision and Rationality. His research interests revolve primarily around the topics of rationality, rational action, epistemology in social sciences, methodology but spin off to reach rethoric of motives, discourse analiysis.